DOCUMENT RESUME

ED 111 681 SE 019 680

AUTHOR Steffe, Leslie P., Ed.
TITLE Research on Mathematical Thinking of Young Children: Six Empirical Studies.
INSTITUTION National Council of Teachers of Mathematics, Inc., Reston, Va.
PUB DATE 75
NOTE 207p.
AVAILABLE FROM National Council of Teachers of Mathematics, 1906 Association Drive, Reston, Virginia 22091 ($3.90, discounts on quantity orders)

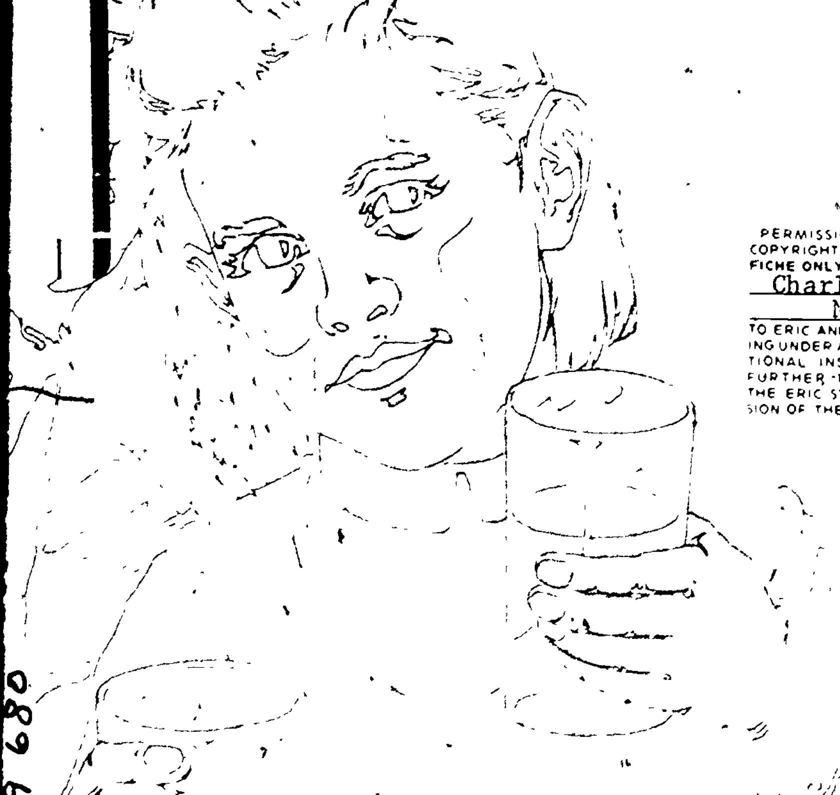
EDRS PRICE MF-$0.76 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Cognitive Development; Concept Formation; Elementary Education; *Elementary School Mathematics; *Learning; Learning Theories; Mathematical Concepts; *Mathematics Education; Primary Grades; *Research; Thought Processes
IDENTIFIERS *Piaget (Jean); Research Reports

ABSTRACT

This volume includes reports of six studies of the thought processes of children aged four through eight. In the first paper Steffe and Smock outline a model for learning and teaching mathematics. Six reports on empirical studies are then presented in five areas of mathematics learning: (1) equivalence and order relations; (2) classification and seriation; (3) interdependence of classification, seriation, and number concepts; (4) Boolean Algebra; and (5) conservation and measurement. In a final chapter, the main findings of these papers are summarized and implications are discussed, with suggestions for further research. (SD)

research on mathematical thinking of young children

SIX EMPIRICAL STUDIES



NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

Research on Mathematical Thinking of Young Children

Research on Mathematical Thinking of Young Children

Six Empirical Studies

Edited by

LESLIE P. STEFFE

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

Library of Congress Cataloging in Publication Data:

Main entry under title:

Research on mathematical
thinking of young children

Includes bibliography.
1. Mathematics--Study and teaching (elementary)--Addresses, essays, lectures. 2. Learning, Psychology of--Addresses, essays, lectures. I. Steffe, Leslie P. II. National Council of Teachers of Mathematics.
QA135.5.R47 372.7 75-22461

Printed in the United States of America

Table of Contents

LESLIE P. STEFFE

Introduction

Acknowledgments

The studies reported in this monograph have been influenced by the work of many professors. In the current zeitgeist in mathematics education, it is easy to underestimate the impact of the foresightedness of a rather small number of professors in mathematics education who did much of their work during the decades circa 1950-1970. Two of the most influential contributors, directly and indirectly, to the works of this volume are Myron Rosskopf, past Professor of Mathematics and Chairman of the Department of Mathematical Education, Teachers College, Columbia University; and Henry Van Engen, Professor Emeritus of Mathematics Education, the University of Wisconsin. Their work stimulated others to examine the formation of operational characteristics of basic mathematical concepts in children of school age. This effort made clear the necessity of the interdisciplinary nature of research in mathematics education. The emphases and substance of research in cognitive-developmental psychology, especially Piagetian, proved most timely and, we hope, most fruitful.

Professor Charles D. Smock, Professor of Developmental Psychology at the University of Georgia, provided many insightful suggestions. His influence, however, transcends the contents and ideas in this monograph; one application of cognitive-developmental theory to mathematics education has been translated into action through his efforts. Others who have directly influenced the contents of this volume are Thomas Romberg, Professor of Mathematics Education, the University of Wisconsin and John LeBlanc, Associate Professor of Mathematics Education, Indiana University. Professor Romberg was dissertation director for Thomas Carpenter, author of Chapter VII, and Professor LeBlanc was thesis advisor for Richard Lesh, author of Chapter V.

Conferences

In 1967, the National Science Foundation, the National Council of Teachers of Mathematics, and the Department of Mathematics Education, University of Georgia, cosponsored a conference on needed research in mathematics education. The conference was general in nature in that it included major papers on needed research in the learning of mathematics, in the teaching of mathematics, and in mathematics curriculum. It was felt by most of the participants that the conference was well-conceived and well-directed for one of the first national conferences on needed research in mathematics education. But, from this conference, it was clear that follow-up efforts aimed at more specific areas of research in mathematics education would be needed--the scope of research in mathematics education is too broad to be dealt with adequately at any one conference.

Subsequently, a conference was held on Piagetian cognitive-development research and mathematical education. This second conference was sponsored jointly by the National Council of Teachers of Mathematics, the Department of Mathematical Education, Teachers College, Columbia University, and the National Science Foundation. The proceedings from this conference consisted of 14 major papers. These papers were successful in identifying the state of the knowledge in the area of cognitive development research in mathematics education and in highlighting promising areas for further research.

It was clear to some of those participating in the conference, however, that the papers were theoretical and experimental in nature and did not, to any great extent, deal with mathematics pedagogy and Piagetian cognitive-development theory. Consequently, a symposium was held on Piagetian cognitive-development research and mathematical laboratories at Northwestern University in 1973 as part of the dedication ceremonies for their new education building. These papers ranged from theoretical expositions of cognitive-development theory to practical expositions of teacher education programs.

These two conferences and one symposium, in addition to theoretical papers, empirical research, and related projects by mathematics educators and psychologists, attest to the exploding interest in the United States in the area of application of cognitive-development research in mathematics education research and development. The papers of these conferences and the symposium are significant contributions to literature in mathematics education. However, the concepts and principles contained in the papers are far from being universally applied (or acknowledged) in educational practice in mathematics education in the schools and colleges of the United States. Such a state of affairs is not necessarily undesirable in view of (1) the lack of formal training in cognitive-development of professional educators in mathematics education and (2) the state of the applied research of cognitive-development to mathematics education. As difficult as it may have been to institute changes in the school mathematics curriculum in the 1950's and 1960's, the mathematical preparation of professional mathematics educators in those two decades far excels psychological preparation of mathematics educators in the 1970's. Even so, there has been a dramatic Piagetian renaissance in mathematics education in the United States over the past ten years. Much work remains to be done, however, if the children in the schools are to realize the potential benefit of recent advances in knowledge derived from cognitive-development theory and research.

On an Outline of a Program of Research

A massive amount of theory and data exists which describes the development of mathematical and scientific concepts in children from the onset of the formation of the permanent object through adolescence--much of that theory and data was generated by the Genevan school within a specific epistemological framework. This theory and data, while it is extremely rich, certainly was not generated by researchers primarily interested in the establishment of scientific pedagogy. As such, it cannot be indiscriminately applied with the hope that, somehow, such application will improve the state of affairs in mathematics education.

Generally, mathematics educators are concerned with the child's learning aspects of various mathematical systems in school mathematics--the natural numbers, the integers, the rational numbers of arithmetic, the rational numbers, the real numbers, polynomials, Euclidean geometry, transformational geometry, linear spaces, matrix theory, and finite systems, to name some. Cognitive-development theory can contribute to an understanding of how it is a child acquires knowledge of these mathematical systems through its description of cognitive operations children acquire and the mechanism through which children acquire them. A mathematics educator cannot stop there, however, because the cognitive operations demanded by mathematical systems may be distinguishable from (but include) the cognitive operations described in cognitive-development psychology. Mathematics educators do not yet know how to utilize the cognitive operations studied in cognitive development psychology in the further acquisition of cognitive operations demanded by the mathematical systems mentioned. In fact, few attempts have been made toward the identification of relationships between the cognitive operations studied in developmental psychology and the cognitive operations demanded by the mathematical systems. For example, (1) if a child is or is not in possession of cognitive operations normally attributed to the grouping structures vis-à-vis Piaget, what does this say about his knowledge or acquisition of the integers, of the rational numbers of arithmetic, or even the rational number system? Or, (2) if a child does or does not possess the proportionality scheme or the INRC Group, what does this say about his knowledge or acquisition of measurement, of the rational numbers, of finite algebraic systems?

Not only, then, is it critical to test for possible relationships between the cognitive systems of the child and knowledge of the systems of mathematics, it is also critical to learn how the mental operations normally attributable to the grouping structures, figurative structures, and the formal operational structures are utilized by the child in the acquisition of mathematical content whose structural properties are not necessarily isomorphic to those genetic structures. All of this information is difficult to acquire, but until the story is told, critical assumptions will have to be made in the application of cognitive-development theory to mathematics education--assumptions which, of course, may not be untenable.

Assuming the validity of certain critical assumptions alluded to above, applications of cognitive-development can be made to mathematical education in which learning-instructional models can be formulated and tested empirically. Such a model may not attain the status of a theory but can be used to both describe and prescribe learning-instructional phenomena concerning mathematics until it proves unusable in terms of the desired objectives and/or learning process.

LESLIE P. STEFFE
CHARLES D. SMOCK

On a Model for Learning and Teaching Mathematics

Piaget (1971, p. 21; 1973a, p. ix) has expressed the convictions that (1) experimental pedagogy must remain distinct from psychology but, yet, utilize psychological principles, and (2) hypotheses derived from psychology must be subjected to empirical verification (or refutation) rather than be accepted only on the basis of deduction. While it would seem unnecessary to restate these two convictions, education has a history of embracing a particular theory or point of view based on overly simplistic deductions and then, when educational practice has shown the transparency of these deductions, abandoning that theory. Piagetian cognitive-development theory may be no exception to this general pattern in education because it is, in the main, the theoretical basis on which the "mathematical laboratory" is built.

The mathematical laboratory, according to Smock (1973) has "remained only loosely or ambiguously defined (p. 1)." If that assertion remains true after substantial attempts are made at an unambiguous definition, then the mathematical laboratory would have to rely on only sloganese or personal testimony for its justification as an instructional approach. Until and unless mathematics educators return to the beginning and ask "how do children learn?" rather than "how do we teach?" the fundamental dimensions of educational and instructional problems facing the mathematics educator will remain unidentified. If we confront this basic question, it is possible to proceed with the task of characterizing what is meant by a "mathematical laboratory." The emphasis, then, is to be placed not on methodology and learning activities but, rather, on the learning characteristics of the child as he acquires particular subject matter and content and methodology constructed for that special purpose.

The preparation of this chapter was supported, in part, by the Mathemagenic Activities Program, Follow Through, C. D. Smock, Director, under Grant No. EG-0-8-522478-4617 (287), Department of HEW, USOE.

The problems of characterizing the mathematical laboratory can best be understood in the context of the issues involved in developing a theory of mathematical *learning* and *instruction*. Bruner points out (1964b) that most theories of learning and development are descriptive rather than prescriptive. While this may be the case, a theory of mathematical instruction for children must be based in the developmental constraints of concept learning, a theory of learning relevant to mathematical concepts, and in mathematics itself if the theory is to have validity. Consequently, rather than focusing on instruction with little regard for learning or on learning with little regard for instruction, analysis of school mathematics requires that learning and instruction be considered simultaneously. Even then, little progress will be possible if mathematics and cognitive development are ignored.

One of the best examples of an analysis and synthesis of cognitive development theory, fundamental mathematical structures, and mathematics instruction for the early school years has been presented by two Russian educational psychologists, El'konin and Davydov (1974). Following Vygotskii, a child's mental development is viewed as being ultimately determined by the content of the knowledge studied. They feel that researchers who study the development of mental operations (notably Piaget) concentrate only on those processes which are maximally independent of specific subject matter. El'konin and Davydov are critical of this approach because it leads to a view that the sources of mental development lie in the individual independently of the specific historical conditions of existence and characterize the child's mind in absolutist terms. On the other hand, they do not ascribe to Bruner's hypothesis that "any subject can be taught effectively in some intellectually honest form to any child at any stage of development." Such an assumption makes reference to abstract forms of teaching fundamentals of any subject to a child of any age, but forms of instruction must be found that are suitable for each specific piece of content *and* given age level.

Piaget (1971, p. 21), however, clearly differentiates experimental pedagogy from psychology. Experimental pedagogy is concerned less with the general and spontaneous characteristics of the child than with their modification through pedagogic processes. Moreover, in commenting on the value of developmental stages in educational sciences, Piaget (1971, p. 171) rejects the notion of inflexible stages characterized by invariant chronological age limits and fixed thought content. As an interactionist, Piaget (1971, pp. 171-73) advocates that the cognitive structural changes that come about through maturation and those that derive from the child's individual experience be considered as separate factors in intellectual development. Also, Piaget (1964a) maintains that mathematical structures can be learned if the structure of interest can be supported by simpler, more elementary structures. Consequently, while the Genevans' work has not been in experimental pedagogy but has dealt with the development of the child independent of particular subject matter curricula, experimental pedagogy does not stand in opposition to cognitive-developmental psychology. Rather, experimental pedagogy is complementary to it, with the potential of contributing knowledge of the developmental processes.

Piaget rejects the notion of inflexible stages characterized by invariable chronological age limits and fixed thought content. Yet, El'konin and Davydov make a central issue out of whether to characterize a given age level in terms of processes for which a developmental period is concluded or in terms of the processes for which that developmental period is beginning. The former is rejected because it would lead to presentation of educational exercises that demand only previously formed intellectual processes for solution and to the further assumption that intellectual development is inviolable and independent of the content and methods of presentation of subject matter. El'konin and Davydov, thus, believe that development of psychological processes underlying the learning of mathematics not only *do not precede* instruction in mathematics but are formed in the process of learning.

El'konin and Davydov's point of view is consistent with the leading role ascribed to instruction by Soviet psychologists (Kilpatrick and Wirszup 1969, p. v). But El'konin and Davydov do not separate methods of instruction from the content of what is to be learned nor from the general cognitive development of children. El'konin and Davydov test experimental curricula based on measurement processes, but their experiments do not provide conclusive evidence that the cognitive development of the children was altered. The experiments shed some light on symbolization capabilities of concrete operational children in highly structured measurement exercises, but little evidence is presented which would lead one to believe these concrete operational children were, as a result of the instruction in measurement, in the formal reasoning stage vis-à-vis Piaget. Thus, the contribution of school learning to cognitive development (in a Piagetian sense) remains to be adequately tested.

The important contribution of El'konin and Davydov is the detailed analysis of fundamental mathematical structures and the explication of the yet unconfirmed hypothesis that cognitive development of children can by altered by school instruction in mathematics. Their analysis of the fundamental mathematical structures led them to the conclusion that the concept of quantity, with its roots in the ordering structures, should be the starting place for school mathematics. The task of analyzing the similarities and differences of the fundamental mathematical (including those mentioned by El'konin and Davydov) and genetic structures is an important first step (Beth and Piaget, 1966).

One important difference is that mathematical structures are the object of reflection by the mathematician but genetic structures are manifested only by the child's behavioral-action structures, which are determined by his assimilation of past experiences. A second major difference is that form is *independent* of the content in the mathematical structure but in genetic structures the form is *inseparable* from content. Finally, in the mathematical structures axioms are the starting point of formal deduction whereas in genetic structures the laws are the rules which the child's deductions obey.

Similarities in the structural types are twofold: (1) operations in the mathematical structures correspond to operations in the genetic structures; and (2) the axioms of the mathematical structures correspond to the "laws of combination" in genetic structures. It is these operational genetic structures which El'konin and Davydov identified as being maximally independent of specific subject matter. The basic content of the genetic structures, however, are classes and relations which, in synthesis, form the basis of cardinal and ordinal number in the logical domain and of quantity and of measurement in the infralogical domain. On the face of it, then, it seems as if El'konin and Davydov's declaration of genetic structures as maximally independent of school mathematics is unjustified. It is true, however, that profound structural differences exist in the genetic structures of concern and certain of the mathematical structures which are used as the basis for the content of school curricula. These structural differences are a central issue in the determination of the applicability of genetic structures to mathematical learning. Mental operations associated with such mathematical structures may or may not be accounted for by the mental operations ascribed to the genetic structures. If the latter case is so, El'konin and Davydov *may be* essentially correct in asserting that the psychological processes are developed concomitant with learning. Piaget (Beth and Piaget, 1966, p. 189), however, proposes that the construction of mathematical entities is an elaboration of the elements of natural thought and the construction of mathematical structures is an enlargement of particular mathematical entities.

The Piagetian hypothesis is, at least, intriguing--it suggests that the individual *constructs* his own mathematics. The testing of the hypothesis, however, is extraordinarily complex due to the structural differences alluded to above. So, two programs of research would seem to be necessary. On the one hand, the logical and mathematical veracity of Piaget's system, concentrating on similarities and differences between genetic and mathematical structures, requires both logical and empirical study. On the other hand, models (albeit preliminary) for learning and instruction of mathematical concepts need to be developed where that development explicitly includes theoretical as well as empirical components. It is this second program of research which is of basic concern in this monograph.

Piagetian theory offers several principles which can be utilized in the construction of such a model (Piaget, 1971, 1973b; Smock, 1970, 1973). First, equilibration theory providea a theoretical model of knowledge acquisition with specification of the factors that regulate acquisition. Second, two distinct levels of cognitive functioning--figurative processes and operative processes--are identified as necessary for understanding learning and development. Third, cognitive capacities determine the effectiveness of training and these cognitive capacities are influenced by four factors that contribute to cognitive development. Fourth, the learning environment must be considered from both the points of view of the genetic structures and the mathematical structures. The implication of some of these principles for the development of an instructional model and for research relevant to that model is presented below.

General Factors Contributing to Cognitive Development

Piaget (1964a) has identified the major factors contributing to the development of cognitive growth of children as including: (1) maturation, (2) experience, (3) language, and (4) equilibration.

Maturation

The proposition that *maturation* is a major determinant of cognitive growth is not, of course, a novel idea. What is new in Piaget's analysis is the explicit inclusion of a maturational *component* as a modulating factor to the experimental and/or experiential contribution to development progress. The constraining role of maturation is supported by, for example, the fact that transitive reasoning seldom has been observed in children below four years of age. The evidence indicates great difficulty in a child's learning transitivity of, for example, "as many as" during the stage of preoperational representation (even when apparently appropriate learning experiences have been encountered). Such evidence cannot be taken as *proof* that maturation is responsible, but it certainly suggests that instrinsic physio-biochemical processes play a prominent part in development of thinking. Subjecting a child to learning "experience" does not appear sufficient to insure he will understand the concept of transitivity.

Experience

Experience in and of itself may not be sufficient to explain conceptual learning of children but no one denies its importance for intellectual growth. But, if experience was sufficient, all one would have to do to "teach" transitivity would be to give the child sufficient exposure—and he would learn. But, unfortunately, it is not that simple.

Physical Experience and Mathematical Experience. Piaget (1964a) has analyzed "experience" into two components: physical and logico-mathematical experience. Imagine, for example, that a child matches objects of Set A one-to-one with objects of Set B through overt (practical) actions; i.e., he places one object from Set A in correspondence with one object from Set B, etc., until all the objects of one (or both) of the sets are exhausted. Then, he takes the objects of Set B and likewise matches them with the objects of another set, C. Does this matching constitute a physical experience or a logico-mathematical experience? It could be either, depending on the cognitive level of the child.

One cannot differentiate between the two types of experiences through observation of the child's overt acts of matching. The crucial determiner of the type of experience is whether the Sets A and C are "related" by the child by virtue of the comparisons of A and B and then, B and C. If the child is not able, through reasoning (mental operations), to determine the relation between A and C, then the experience gained through overt matching of the objects of A and B and B and C was, by definition, physical in nature. The relation between the Sets A and B in this case was a function of the physical arrangement of the objects and would not exist for that child in the absence of perceptual input. That is, the *relation* remains *external to the child* and thus is destroyed upon rearrangement of the objects of the sets. When the two sets of objects are in a state for physical comparison, the child definitely obtains knowledge about the objects—either they match or they don't—but, for that knowledge to be mathematical in nature, the *relation* must be conserved by the child when the objects are moved to new states and the child must be able to engage in reasoning involving the properties of the relations that go beyond the perceptually or graphically "givens."

The distinction between a physical experience and a mathematical experience is essential to the understanding of the growth of mathematical concepts. Knowledge based on physical experience alone is knowledge of static states of affairs and, if a child is wrong, it is easy to demonstrate that to him. However, knowledge derivable from mathematical experience is another matter; if a child is wrong, it is difficult, if not impossible, to convincingly demonstrate, or even to get the child to accept verbal explanation of the correct answer. For example, if a child fails to *align* the two endpoints when comparing length of sticks, it is quite easy to correct the mistake. If, however, he fails to display transitive reasoning in a task, one or two examples is not likely to teach him the concept.

Physical knowledge, then, is the construction of the invariants relevant to the properties of objects (i.e., states) and is based on "experience" through direct contact with objects through one or more of the five senses. For example, one may touch something and it is hard, cold, hot, soft, supple, etc. Or, one may see something—an object is red, a diamond cutting glass, the shape of a banana, etc. The source of mathematical experience, however, is assumed to be the abstractions from coordination of actions vis-à-vis object; i.e., transformation of the "states" associated with series of discreet physical experiences. The critical difference is that the mathematical knowledge gained demands that a pair (or set) of physical objects not be defined by the temporal-spatial (perceptual) similarities, but rather by the invariant relations among or between objects. Overt (perceptual) actions alone are not sufficient for mathematical experience. As already noted, a child may match the objects of two equivalent collections and the knowledge gained from the actions and perceptual consequences may be no more than physical knowledge. Often cited as an example of mathematical experience is the realization by a child that it makes no difference *how* you count a collection of objects--you get the same number. Such knowledge is gained only through counting the collection in at least two or more ways; i.e., coordination of mental, as well as practical, actions.

Linguistic Transmission

Language, the third factor in the growth of mathematical concepts, is considered a part of the experience of the child, but deserves special consideration because of its special quality and status within the total realm of experience. Information contained in a verbal communication will not, necessarily, increase a child's understanding of a mathematical concept. Bailey (1973), for example, presented a transitivity problem to 40 first graders, 40 second graders, and 40 third graders who were in the top two-thirds of their class according to teachers' judgments. Each child was presented a transitivity of length problem. Those children who did not solve the problem were told the correct relations between the two sticks. For example, if a child took a stick B and first compared it with C, and then with A, and found that A and B were the same length and B and C were the same length, but could not infer the relation between A and C, he was told A and C were the *same length*. After the verbal instructions, the child was asked to explain why A was as long as C. Of 24 children who could not infer the correct relation between A and C, only five would give a satisfactory explanation of why A and C were of the same length *after* the verbal instructions.

It would appear that for children in the stage of preoperational representation (or even in the transitional stage), any attempt to teach mathematics concepts only through verbal or symbolic means will be unsuccessful. But, because words and symbols are an important part of mathematics, their specific functions must be clarified. Until then, a carefully arranged interplay between the spoken words which symbolizes a mathematical concept and the sets of actions performed in the process of constructing a tangible representation of the concept should be maintained. In short, a mathematical vocabulary should be developed *during* the course of activities used to explicate and provide embodiments for a concept to be learned. The particular blend, of course, will be determined by the specific activity and characteristics of the child.

Equilibration and Learning

Of the four factors which contribute to the growth of mathematical concepts, Piaget considers equilibration to be most fundamental. Equilibration is a self-regulatory mechanism that balances the invariant biological adaptation processes of assimilation and accommodation. Assimilation refers to the process by which novel events are integrated into the existing mental structures. The complementary process of accommodation concerns the alteration of mental structure under the pressure of this new information.

Learning, in this context, refers to the process by which new information is assimilated into available cognitive structures *and* to the modification of those structures (accommodation). Mathematical learning, then, appears to be more than association of stimulus and response. The *association* of 6 with (2 x 3) is important and no one doubts it can be conditioned (or memorized) using appropriate instructional strategies. Teaching which assumes a stimulus-response view of learning runs the risk of promoting physical knowledge and not mathematical (or operational) knowledge. However, because learning is generally thought to be provoked by situations external to the learner, it is necessary to analyze the levels of learning relevant to particular mathematical concepts or structures.

If our assumption is correct, there should be a differential emphasis on assimilation and accommodative activity depending on the level of understanding of a particular concept. Assimilative task structures would emphasize "play" and/or self (child) regulated activities until the child's behavior indicates the essential elements of a concept have been assimilated. Then, more task situations (including modeling and verbal exploration), as well as situations designed to utilize and generalize the relevant concepts to new situations, should be introduced. The appropriate balancing and sequencing of the assimilative-accommodative activities (practical and mental) requires considerable theoretical and observational skills of the instructor. At the same time, the basic ideas and associated techniques can be identified and used as guidelines for a Piagetian type of learning environment.

Levels of Mathematical Concepts. Relations, classes and number take a relatively long period to develop as operational concepts in the child, appearing perhaps as late as nine to ten years of age. The stages of cognitive development in Piagetian theory are identified as sensorimotor, intuitive preoperational, concrete operational, and formal operational with the movement from one stage to the next being clearly marked by "transitional" phase characteristics. Our assumption is that mathematical concepts go through similar "stages" (called levels) as the child "learns" a concept; e.g., in the case of relations, either children have little or no knowledge of relations, or they are able to engage in reasoning involving the properties of the relations, or they are oscillators —at times, in restricted situations, they appear as if they are able to reason involving relations, but that reasoning is limited and can be extinguished quite easily. Finally, as the assimilation-accomodation activities are "balanced," the child now "feels" he "understands" and insists on the "logical necessity" of the concept.

Learning-Instructional Phases for Mathematical Concepts

Phase I: Exploration

Equilibration, the balancing of assimilatory and accommodatory activity, is a useful theoretical construct to help determine the criteria for learning activities in mathematical instruction. Exploratory activities have been identified by Piaget as representing a preponderance of assimilatory activity; i.e., modification of the environment to match the existing cognitive structures. As such, exploration is conceived as an essential first step in mathematical instruction and learning. This first phase corresponds to the first level of mathematical concepts identified; i.e., to essentially "no concept" and to the emergence of the second level (rudiments of a concept). At these two levels of mathematical concept development, emphasis is on the *constructive* thinking by the child. It is a period of concept formation and not analysis. Exploratory activities can vary along two dimensions: the *type and structure of the material* and the *degree of direction* given to the child. However, it is important to keep in mind that the child needs to structure (assimilate) the activities but in a direction relevant to the particular concept learning desired.

Multiple Embodiment Principle. In order to illustrate the above principles using particular concepts, imagine that one-to-one correspondence is the concept of interest. The "no concept" phase of one-to-one correspondence corresponds exactly to the preoperational stage of development. Children who display lack of one-to-one correspondence first must be allowed to engage in undirected exploratory activities, using physical objects that later will be used in more directed activities. For example, the child may be given assortments of beads, bird cutouts, blocks, discs, animal cutouts, toy animals, toy cowboys, toy soldiers, etc., and allowed free play time with these materials. Most children will attempt to place cowboys and Indians on horses, dress dolls, stack dishes, stack blocks, string beads, categorize animals or bird cutouts, align soldiers in rows and give each guns, etc. Further, this type of *practical play* is extended *by the child* into symbolic play; i.e., such as waging wars, keeping house, setting tables, etc. The length of time and number of free play activities which should be encouraged is determined by the frequency of "transitional indication" in the child's behavior and two additional principles: (1) *multiple-embodiment* and (2) *mathematical variability* (Dienes, 1971).

The first (multiple-embodiment) states simply that in play activities, the child should use as many different material sets as appropriate so long as each material set is conducive to construction (by the child) of the concept. Consider a particular free play activity where the preoperational child places cowboys and Indians on horses. Through this assimilatory activity, the child can gain the physical knowledge that indeed the cowboys and Indians fit on the horses. The situation is set for an adult to create a disequilibrium for the child by asking the child if there are enough cowboys and Indians so each horse would have a rider. To find out, the child has to change his practical activity to answer that question. Now if the child correctly achieves that task, the next step is to introduce new materials but maintain the original goal (one-to-one matching); e.g., are there enough dresses so one could be put on each doll?

If the child does not initiate specified goal-directed activities following suggestions, the adult may then demonstrate that there are enough cowboys and Indians so each horse has a rider. Such *imitation* learning (accommodatory activities) provides the conditions to employ the principle of multiple-embodiment for subsequent learning tasks and even, at times, imitative behavior across task situations. The adult must, always, be sensitive to the type of knowledge the child is acquiring in the imitative activities; i.e., knowledge acquisition under imitative conditions has a high probability of being at the level of "physical knowledge" if children are in the preoperational stage. By remaining sensitive to this, the adult will avoid expecting the child to build an understanding of higher-order concepts prior to acquiring the necessary prerequisite concepts.

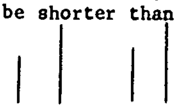
Mathematical Variability Principle. It was pointed out above that free play activities can, through appropriate intervention, be changed into more directed learning activity for children. The second dimension of task activity that can be modified by the adult involves the application of the mathematical variability principle (Dienes, 1971). In contrast to the multiple embodiment principle where the mathematical content is held constant and the materials varied, the mathematical variability principle varies the mathematical content. For example, in case of one-to-one correspondence, the relation being considered can either be changed to a new relational category altogether (e.g., length relations or family relations) or to a relation within the category of matching relations (i.e., more than, fewer than, as many as). Both of these types of variations can, if appropriately used, create cognitive conflict to be resolved by the child.

In summary, it has been pointed out that "play" can vary from free play to directed play, where the directed play is a natural extension of free play. Further, the principles of multiple embodiment and mathematical variability provide guidelines for setting the stage for transition from the level of physical experience to mathematical experience. However, dramatic short-term success in teaching mathematical concepts to preoperational children is not to be expected. More success can be expected with children in the transitional stages (which corresponds here to the second level of mathematical concepts), but again, the short-term success will undoubtedly be modest. During the developmental phase of preoperational representation, it is advocated that the *learning-instructional* phase conditions be held relatively constant, i.e., use the explorations with variations based on the mathematical variability and multiple embodiment principles.

Phase II: Abstraction and Representation

The second learning-instructional phase, that of abstraction and representation, is based in part on the distinction between physical experience and logico-mathematical experience. Abstraction from properties of objects (physical experience) are called simple abstractions; e.g., hardness, sharpness, etc. Reflective abstraction (logico-mathematical experience) involves abstraction from the *actions performed on* (or with) objects *or* representations of those actions. Reflective abstraction is represented by the case of a child counting a string of beads from one end, then from the other, and realizing that the number of beads is independent of the manner of counting them; i.e., which is selected first, which second, which third, etc. The beads are there, but the knowledge gained had to do with the actions with the beads and the capability of representing and reviewing the actions. Or a child may pair elements from two sets until one is exhausted prior to the other; re-pairing the elements in a different way provides the conditions necessary to make the abstraction that no matter how the pairing is done, the one set will always contain more elements than the second set. It is clear that a child may be "playing" but still be engaged in reflective abstraction. Thus, the teacher need not be restricted to "play" activities as long as the child reveals capacity to make the higher level abstraction.

The child may engage in reflective abstraction but not achieve stable representation of newly gained knowledge. Representations may be figurative (e.g., images) derived from drawings or perception of a collection of symbols, etc., and not available to the child at a later time. For example, if a child compares a green stick with a red stick and finds the green stick to be shorter than the red stick, any one of



G < R, G < R, or G R

might be used as a static, immediate representation. There, reflective abstraction and representation together contribute to a higher level of concept formation than expected in the first learning-instructional phase in that rudiments of mathematical concepts are present and can be utilized in a limited way (2nd level of concept formation).

Phase III: Formalization and Interpretation

The learning-instructional phase of formalization and interpretation completes the proposed learning cycle for mathematical concepts. The mathematical concept base-ten numeration system will be used to illustrate the three learning-instructional phases, with emphases on formalization and interpretation (Phase III). The concept was selected because no data are presently available that show the numeration systems are part of the natural cognitive development of children.

We assume the child is at the concrete stage of operations and the operational structures related to classification and relations are available to him. It must be noted that because a child is at this developmental stage does not mean he knows base ten numeration nor that figurative representation of the concept has been achieved. What use the child is able to make with his knowledge of classes, relations, and number may be in the absence of formalization of numeration concepts. Therefore, the only prerequisites required are the completion of a learning cycle concerning the ordering of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9; ability to write them from memory, and to do simple addition problems.

Any natural number can be written in expanded notation. Consequently, the coefficients, the base, and the exponents can all be allowed to vary in employing the mathematical variability principle. Usually, the base is held constant (base ten is used) and only the exponents and coefficients are allowed to vary.

Imagine that a small group of children are given a collection of various assortments of materials, such as geometrical shapes, checkers, dried beans, etc., and are allowed to engage in free play with the materials, building whatever they wish—castles, houses, roads, forts, etc. The first type of direction which could be given to the children is to find "how many piles, with a certain number in each pile, can be made from the objects." The mathematical variability principle should be employed by varying the number in each pile or the total number of objects in each collection. The multiple embodiment principle should be used by changing the type of objects (thus setting a new problem each time) or the type of collection to be formed. The essential aspect is that a collection of objects always can be partitioned into *subcollections* with the *same number* in each and *one other subcollection* with relatively *fewer objects* in it; e.g., a collection of twenty-six objects can be partitioned into four subcollections with six per subcollection *and* two objects in the nonequal set.

One of the first bits of logico-mathematical knowledge the children generally acquire is the sameness of the *number* of objects in the total collection before and after the partitioning process. That is, the child discovers, through his actions, that a pile of objects can always be put back the way it was before the partition (reversibility) and that no objects were added or subtracted; therefore, the number of objects before and after piling is the "same" even though the child *does not* know how many things there are.

Specifically, if a child starts with some objects, makes three piles with six per pile, and one pile of four, the child should know that the total number of objects in the original pile is the same as the number of objects in three piles of six and one pile of four, without knowing there

objects ⟶ partitioning action ⟶ pile of 6
pile of 6
pile of 6
pile of 4

are 22 objects. The mathematical variability principle (varying the number of objects in each pile) should help the child to the realization that no matter how many are in each pile, the total number in *all* the piles is equal to the total number of objects. When a child discovers this, he has made a transition from the first learning-instructional phase (exploration) to the second learning-instructional phase (abstraction and representation).

At this point, the operational basis for constructing the concept of a numeration system has been laid. The goal of the second learning-instructional phase is to have the child construct a notational system and construct the place-value concept. Generally, the two-digit numerals are worked on at different age levels than are the three-digit numerals, which in turn are worked on at different age levels than are the four-digit and higher-digit numerals (the generalization of place value).

After the children first enter the second learning-instructional phase they are able to partition a collection into subcollections and know that the number of objects in the original collection is the same as that in all the subcollections. Capitalizing on this knowledge and the ability of the children to engage in rational counting, place-value concepts may be developed. However, until the children can represent any collection of *a* tens and *b* ones (where *a* and *b* are digits) as "ab," they are not ready to learn the number names for the two-digit numbers and order the numbers from 0 to 100. The number names and the order relation are included in the next learning-instructional phase because the knowledge gained to this point is to be systemized by the children. The basis for the learning which is to take place has been laid in counting out piles of ten. However, the main goal of the formalization-interpretation phase is to, again, systemize the whole numbers from 0 to 100 using the number names. Formalization takes place in the sense that a notational system is developed and organized by the child. The organization of the notational system is based on the abstraction and representation accomplished at the second phase and on the new element of an *order relation*. The order relation is an essential part of the third learning-instructional phase for the concept of numeration. Without it, the third phase would have little meaning. The order relation, however, is based on one-to-one correspondence, so that preliminary cycles will have to have been completed with regard to one-to-one correspondence and number.

Some Problems

The foregoing model for learning and teaching mathematical concepts needs critical examination in view of other theoretical constructs in an attempt to build the best first approximation of a model possible. That refinements are possible is easily recognized, as a host of theoretical constructs exist which have not yet been integrated into the model. Some of these constructs are figurations and operations, the concept of decentering, and egocentrism of the child. Moreover, a great deal of difference may exist between the model's applicability to mathematical concepts shown to be part of the general development acquisition and those not. Further, no attempt has yet been made to apply the model to geometrical and spatial concepts. It may well be that a model for learning concepts in those areas is greatly different from a model for learning numerical concepts. However, reason does exist that the above model is applicable to geometrical and spatial concepts as Piaget views the grouping structure as a basic structure of mental operations in the infralogical domain. In any case, it is clear that theoretical refinement of the posited model is necessary and that the model needs to be studied across different mathematical concept areas for applicability. Certainly, there are no a priori reasons that a single model is sufficient to account for the learning of disparate mathematical concepts.

The theoretical problems associated with the model should not preclude empirical study of the validity of certain crucial implications of the model. Especially important are such factors as:

1. Are the four factors contributing to the development of certain mathematical concepts critical to the learning of those concepts which have not been shown to be part of the natural course of intellectual development?

2. Can those mathematical concepts not (yet) shown to be developmental phenomena be ordered along the three levels comparable to the concepts known to be part of the developmental process?

3. What is the validity of the learning-instructional strategy involving the mathematical variability and the multiple embodiment principle, in the context of sequencing learning conditions according to the exploration, abstraction-representation, and formalization-interpretation phases?

4. What is validity of learning-instructional phases? Does a structural integration take place only after appropriate overt actions are internalized through reflective abstraction? That is, is the phase of formalization and interpretation identifiable distinct (psychologically) from the phase of abstraction and representation? Reflective abstraction, as a theoretical construct, is appealing, but does it have psychological credibility in mathematics learning?

The experimental studies reported in the following chapters offer a preliminary test of the notion of reflective abstraction in learning of classes and relations by children who had not yet consolidated concrete operational structures. The set of principles inherent in the model described above together with analysis of the relevant genetic structures and mathematical structures were used as guides for the design of the learning materials. No attempt was made to isolate one or more of the underlying factors which may have contributed to resultant learning. The preliminary model was applied in its totality in all of the training studies so that complex interactions of principles in the model were not explored in the learning of selected aspects of classes and relations.

The learning material utilized in the experimental studies did not progress beyond the abstraction and representation phase, except, of course, for cases in which children themselves went into the third phase. The initial learning-instructional tasks were intended to operationally define the concepts for the children. These learning-instructional tasks were written within the constraints of the learning-instructional phase, Exploration. The children were allowed time to engage in free play activities, but direction from the experimenters was included in a highly controlled context to insure that each of the children engaged in overt activities necessary to operationally define the concepts. It was not expected that the children would progress beyond physical knowledge in these initial tasks. The multiple embodiment principle and the mathematical variability principle were employed in such a way that each material set of interest and each concept of interest were included in these initial tasks. A carefully designed schedule for the introduction of terminology was followed throughout the learning materials.

Other learning-instructional tasks were constructed intending to maximize the possibility of children engaging in logical-mathematical experience. These tasks differed from the initial tasks in that they included operations on, or properties of, the concepts involved (e.g., transitivity, asymmetry, class intersection). Hindsight and foresight (anticipation) activities were utilized in the design of these higher-order tasks. No attempts were made to include figural representations or written representations of the higher-order tasks. The tasks always included manipulatable objects. Internal representations of tasks (images or verbal thought) or spoken language were encouraged, but no controls were included to maximize such representation except in the case of terminology developed to communicate the concept elements (e.g., "pair," "partner," "more than"). The time devoted to the two learning-instructional phases was held constant.

In the study by Steffe and Carey (Chapter II) the mathematical structures of equivalence and order relations were used as mathematical models in the construction of the learning materials. The content of the relational structures was length relations defined for open curves of finite length. In the experiment, each child was used as his own control so that information was available prior to the experiment on conservation and transitivity of length relations. The relationship of relational structures to genetic structures has been discussed elsewhere (Steffe, 1973)--similarities and differences pointed out earlier in this chapter were kept in mind, where the similarities were emphasized owing to the status of relations in cognitive development.

The study by Owens (Chapter III) was also concerned with the reflective abstraction and representation of relational structures. Owens, however, employed two relational categories, matching relations and length relations, where each category included equivalence and order relations. Owen's test of whether reflective abstraction took place in the children demanded that the children be able to apply learned properties of matching relations to length relations, where the length relations were only operationally defined. The learning materials, just as in the Steffe and Carey study, employed the foregoing posited model. Owens not only demanded the learned relational properties be transferred to a different relational category in a test for possible reflective abstraction, but he also administered a problem to the children which demanded that transitivity of matching relations be employed in its solution.

In the study conducted by Martin L. Johnson (Chapter IV), rather than attempting to induce properties of equivalence and order relations property by property, children were immersed in total seriation tasks. This decision was predicated on the theory that relational properties emerge as a result of a total scheme of classification or seriation rather than the other way around. It was felt that reflective abstraction would be given the maximal opportunity to operate in a relatively short period of time (16 instructional days out of 22 consecutive instructional days) with content shown to be developmental phenomena.

Lesh, in the study reported in Chapter V, did not begin with mathematical structures in his learning program but, rather, conducted a pilot investigation using the genetic structures as his starting point. On the basis of the genetic structures, Lesh generated three sequences of tasks, one dealing with seriation, one with number, and one with classification. The tasks were subjected to empirical validation as to their sequential nature and difficulty. A transfer of learning experiment was then carried out where training was given on seriation and classification and transfer to the number tasks tested. The preliminary model presented above was utilized in the construction of the learning materials for the transfer of learning experiment.

A fundamental issue raised by Lesh's experiment was mentioned in discussion of the El'Konin-Davydov experiment. The mental operations Lesh worked with were taken from the genetic structures. Because Lesh was dealing with such fundamental concepts, it was possible for him to proceed as he did. However, Piagetian theory has less to offer concerning possible genetic structures underlying more advanced concepts, and even if mathematical structures are considered, mental operations underpinning these structures may be different from those underpinning genetic structures. In that the results of Lesh's training study are positive, more experiments need to be conducted designed to shed light on mental operations underlying the child's concept of number.

The study by David C. Johnson (Chapter VI) was concerned basically with educational technology. No attempt was made to experimentally determine basic mental operations underlying the child's concept of number. Rather, it was assumed that Piagetian theory was essentially correct as it is concerned with relations and classification. The experimenter was explicitly aware of structural differences in mathematical structures and genetic structures dealing with classes and relations, but he emphasized the similarities rather than the differences.

Of the experiments reported in this monograph, David C. Johnson's experiment provides the best test of the construct "reflective abstraction" in the case of learning concepts shown to be developmental phenomena. In the learning material, children were given definite opportunities to engage in mathematical as well as physical experiences. The transfer tests all demanded mathematical knowledge for successful completion, whereas the achievement measures demanded only physical knowledge. So, a test was possible of the amount of mathematical knowledge the learning material produced in the children.

The study conducted by Carpenter (Chapter VII) was concerned with development of mental operations regarding measurement. The study was not experimental in nature but was related to the previous studies in the monograph by virtue of the content of the tasks, the variables controlled, and its developmental nature. In particular, Carpenter tested the degree to which young children possess the logical structures necessary to assimilate and apply information from measurement processes and attempted to identify some factors involved in the development of measurement concepts. The study provides, within a scope limited by the tasks and factors studied, baseline data for designing experiments relevant to specific questions implied by the model.

LESLIE P. STEFFE
RUSSELL L. CAREY

Learning of Equivalence and Order Relations by Four- and Five-Year-Old Children

Elkind (1967) has categorized Piaget's conservation problems into two categories, conservation of identity and conservation of equivalence.

> *Regardless of the content of these problems, they routinely involve presenting the subject with a variable (V) and a standard (S) stimulus that are initially equivalent in both the perceptual and quantitative sense. The subject is then asked to make a judgment regarding their quantitative equivalence. Once the judgment is made, the variable stimulus is subjected to a transformation, V → V', which alters the perceptual but not the quantitative equivalence between the variable and standard. After completion of the transformation, the subject is asked to judge the quantitative equivalence between the standard and the transformed variable (p. 16).*

In the above conceptualization, a judgment of conservation may be relative to conservation of a quantitative relation or to the identity of V and V'. Even though the possibility of two judgments exists, "It is probably true, nonetheless, that from the point of view of the subject, the conservation of identity is a necessary condition for the conservation of equivalence (Elkind, 1967, p. 17)."

Aspects of conservation exist which are not completely clarified by Elkind's characterization. For example, consider the relation "as many as." If the elements of a set A are in one-to-one correspondence with the elements of a set B, the A has as many elements as B, and vice versa (denoted by A ~B). In a conservation problem involving "~", if the child is asked to make a "quantitative judgment," one must be assured that the child associates at least a one-to-one correspondence with the phrase "as many as." That is, one must be assured that a conservation problem is not a test of terminology. The establishment of the initial comparison is also basic to conservation of length relations between two objects. One may take the point of view,

---

This paper is based on Research Paper Number 17 of the Research and Development Center for Educational Stimulation. University of Georgia, Athens, Georgia (Carey & Steffe, 1968).

moreover, that even though a child may point to the longer of two sticks, he may be basing his judgment on two endpoints only without regard to the relative position of the remaining two endpoints. In this case, one should not be willing to accept that he perceives the initial relation. Clearly, a comprehension of relational terms is a prerequisite to problems in conservation of the relation. The phrase "the same length as" has a quite different referent than does "as many as." While both are equivalence relations they still are different relations. Thus, there seems to be no reason to believe that the ability to conserve one of the two relations implies the ability to conserve the other. Smedslund (1964), in a study of concrete reasoning, observed that 31 children failed one of the two conservation problems involving "same as" and "longer than" while 32 failed both and 97 passed both, which supports the contention that the ability to conserve a particular relation does not imply an ability to conserve another. Moreover, in a conservation problem, the initial relation need not be an equivalence relation. It may be, in fact, an order relation (e.g., "fewer than").

Whether the initial judgment in a conservation problem always involves a judgment of quantitative equivalence is not completely clear. For, if A and B are curves of finite length, then A is the same length as B if and only if $L(A)=L(B)$, where $L(A)$ is a number denoting the length of A, and $L(B)$ is a number denoting the length of B. If $T(B)$ is a transformation of B which is length preserving, then $L(B)=L(T(B))$ implies that A is the same length as $T(B)$. If children cannot associate a number with A and B, then there is no reason to believe that "the same length as" has any quantitative meaning for them. Therefore, under these conditions, there would be no reason to expect children to conserve a quantitative equivalence between A and B. It is entirely reasonable to expect children not to be able to associate a number with a segment but yet be capable of establishing a relation not necessarily involving number between two or more segments, for Piaget, Inhelder, and Szeminska (1960) make a sharp distinction between "qualitative" and "operational" measurement.

> *Qualitative measuring . . . which consists in transitive congruence differs from a true metrical system in that the latter involves changes of position among the subdivisions of a middle term in a metrical system . . . . whereas in qualitative measuring, one object in its entirety is applied to another (p. 60).*

> *While conservation, and hence qualitative transitivity, are achieved at a mean age of 7 1/2, measurement in its operational form . . . is only achieved at about 8 or 8 1/2 (p. 126).*

Before presenting length relations to children below six years of age, it seems necessary, then, to define the relations on a basis that does not assume number. Such a definition follows. Let A, B, and C be segments. A is the same length as B if, and only if, when segments (or their transforms) lie on a line in such a way that two endpoints coincide (left or right), the two remaining endpoints coincide. A is longer than B if, and only if, the remaining endpoint of B coincides with a point between the endpoints of A. Also, in this case, B is shorter than A.

The above definitions are acceptable from a mathematical point of view as the length of a curve is the least upper bound of the lengths of all inscribed polygons. Intuitively, then, one could think of the length of a curve as the length of a line segment where, of course, the lengths are identical. It is essential to note that in the definitions given, children do not have to assign numbers to segments through measurement processes. The definitions are given entirely in terms of a line, the endpoints of curves, betweenness for points, and coincident points on a line and are consistent with Piaget's (1964a) view that "learning is possible in the case of . . . logical-mathematical structures, but on one condition--that is, the structure you want to teach . . . can be supported by simpler, more elementary, logical-mathematical structures" (p. 16). The relations "same length as," "longer than," and "shorter than," as defined, and their properties are more elementary and logically precede measurement. The definitions given above are the results of an attempt, on the part of the investigators, to define the relations in as simple a manner as possible but in such a way that they are still mathematically acceptable.

The relations need not be presented to children by the use of words alone as they may be defined operationally, i.e., defined by physical operations with concrete objects. The physical operations eventually need to be performed by the child himself, because central to Piaget's theory is the fact that the child is active; he gains knowledge through his own actions.

Operationally, then, for a child to find a relation between two "rods," say rod A and rod B, he must place the rods side by side and align two of the endpoints. The relative extension of the two remaining endpoints then determines the relation(s). If rod A is in fact shorter than rod B the child, upon placing A by B, can determine that fact. Through an equivalent action or the same action the child also can determine that B is longer than A. It is through the coordination of these actions that logical-mathematical structures evolve for the child as "coordination of actions before the stage of operations needs to be supported by concrete material. Later, this coordination of actions leads to the logical-mathematical structures" (Piaget, 1964a, p. 12).

If a child establishes a relation between two curves in accordance with the operational definition given, then to conserve the relation, the child must realize that the relation obtains regardless of any length-preserving transformation on one or both of the curves. In other terms the child must realize that, after such a transformation, if the curves are moved back side by side as in the original state, the ends will be still in the same relative positions. Viewed in this manner, the conservation of the relation is essential for the transitive property. Take the example of a child who is presented with two fixed line segments, say, of the same length but not obviously so, and a third segment the same length as the first two and then questioned about the relative lengths of the two fixed segments (which he must not overtly compare). The child must realize that once he has established a relation between the lengths of two segments, the relation obtains regardless of the proximity of the segments.

If conservation of identity is viewed in terms of quantity, then there is little logical reason to expect "conservation of identity" to be involved in conservation of length relations as discussed above. While the length of an object is a number assigned to the object, it is not necessary for a child (or an adult) to know the lengths of two segments in order to establish a length relation between them, on an operational basis. The only aspect of identity involved, then, would seem to be that the child would have to affirm that the segment is the same segment no matter how it is moved around.

The length relations as operationally defined and mathematically defined do have certain properties. The relation "the same length as" is reflexive, symmetric, and transitive and the relations "longer than" and "shorter than" are nonreflexive, asymmetric, and transitive. Because for any curve A, A is the same length as itself and not longer or shorter, it would seem that, at least in a relational sense, what sometimes passes for a test of conservation of identity is no more than a test of the reflexive and nonreflexive properties. Although it is not true for equivalence relations in general, if it is assumed that for each curve A there is a curve B such that A is the same length as B, the reflexive property of "the same length as" can be deduced from the symmetric and transitive properties. For if "~" denotes "the same length as," A~B and B~A implies A~A for each curve A. This logical interdependency gives little hope for conservation of identity to be necessary for transitivity.

In addition to the properties of the relations, the following statements are logical consequences of the definitions of the relations given. (1) A shorter (longer) than B is equivalent to B longer (shorter) than A; (2) if A is the same length as B then A is not shorter (longer) than B; (3) if A is shorter (longer) than B then A is not longer (shorter) than B.

Questions and Hypotheses of the Study

Few data exist concerning proficiency levels of four-and five-year-old children in establishing length relations in accordance with the operational definition given. For open curves, the operational definition is extended as follows. To establish a length relation between two open curves A and B, a child must (1) place each curve on a line in such a way that two endpoints (left or right) coincide, (2) compare the relative position of the two remaining endpoints, and then (3) on the basis of (1) and (2), determine what relation holds. Given that a child is able to establish a length relation between two curves, it is hypothesized that he will be able to conserve the relation established. This hypothesis is advanced due to the fact that for a child to establish a length relation, he must attend to relative positions of the endpoints of the curves as well as ensure that the curves are on a line. Even though the actions of establishing length relations between curves are physical actions, for a child to carry the physical actions out spontaneously, it would seem that the child must be operational in a Piagetian sense. If such is the case, a multitude of potential physical actions would be possible, which should include use of the properties and logical consequences of the relations. A second hypothesis is then advanced. If a child is able to conserve length relations established, it is then hypothesized that he would be able to use properties and logical consequences of the relations. It must be made explicit that it is not hypothesized that use of the reflexive and nonreflexive properties precedes conservation of length relations. Rather, it is hypothesized that conservation of length relations is necessary for use of the reflexive and nonreflexive properties as well as the asymmetric and transitive properties. It is further hypothesized that use of relational properties and consequences will be consistent with logical interdependencies of those properties.

All of the above hypotheses are advanced only for children who have not received formal instruction on relations. Because any hypothesis advanced for children who have received specific instruction must by necessity be at least provisionally instruction-specific, a list of specific questions is presented rather than hypotheses. The first list is a question asked for four- and five-year-old children who have not been engaged in formal instruction on establishing length relations; the second list is questions asked for four- and five-year-old children who have been engaged in formal instruction only on establishing length relations; and the third list is questions asked for four- and five-year-old children who have been engaged in formal instruction on establishing length relations, conserving length relations, and using properties and consequences of length relations.

Question asked for four- and five-year-old children with no formal instruction on length relations.

1. What is the proficiency level of children when spontaneously establishing length relations between two curves?

Questions asked for four- and five-year-old children with formal instruction only on establishing length relations.

2. What is the proficiency level of children when establishing length relations between two curves?
3. Does formal instruction only on establishing length relations improve the proficiency level of children when establishing length relations?
4. Are children able to conserve length relations when the asymmetric property and logical consequences are involved as well as when they are not involved?
5. Are children able to use the reflexive and nonreflexive properties?
6. Are children able to use the transitive property of length relations?
7. Is the ability to use the reflexive and nonreflexive properties necessary (or sufficient) for children to conserve relations?
8. Is the ability to use the reflexive and nonreflexive properties necessary (or sufficient) for children to use the transitive property of length relations?
9. Is the ability to conserve length relations necessary (or sufficient) for children to use the transitive property of length relations?

Questions asked for four- and five-year-olds with formal instruction on establishing length relations, conserving length relations, and using properties and consequences of length relations.

Questions (5)-(9) are repeated here as questions (10)-(14).

15. Does formal instruction on conserving length relations; on the reflexive, nonreflexive, and asymmetric properties; and on consequences of length relations improve the ability to (1) use the reflexive and nonreflexive property, (2) conserve length relations, and (3) use the transitive property of length relations?

Procedure

Subjects

The subjects were 20 four-year-old and 34 five-year-old children.

At the initiation of the study, the range of ages was 47-57 months for the group considered as four-year-olds and 59-69 months for the group considered as five-year-olds. The children were in three self-contained classrooms with some of both age groups in each room. The verbal maturity and intelligence of the children were measured by the Peabody Picture Vocabulary Test and Stanford Binet Intelligence Scale, Form L-M (Table 1). The mean for the intelligence of the four-year-olds was modestly higher than that for the five-year-olds.

Table 1
Verbal Maturity and Intelligence

	Verbal Maturity		*Intelligence*	
Age Group	Range	Mean	Range	Mean
4	83-119	102.6	98-145	119.6
5	55-120	97.7	81-130	109.1

According to the Hollingshead Two Factor Index of Social Position, the social classes of the children ranged from I (high) to V (Table 2). Category III for each age group contained the greatest number of children.

Table 2
Social Classes by Age Group

Age Group \ Social Class	I	II	III	IV	V
4	3	4	9	4	0
5	3	8	13	6	4

*Instructional Sequence and Measuring Instruments**

Instructional sequences. Three instructional sequences were constructed especially for the study. Instructional Sequence I was designed to develop the ability of children to establish a length relation between two curves; Instructional Sequence II was designed to develop the ability of children to use the reflexive and nonreflexive properties; and Instructional Sequence III was designed to develop the ability of children to conserve length relations and use the asymmetric property and logical consequences. The following principles were employed in the design of the sequences.

1. Mathematical concepts are not implicit in a set of physical materials. A child gains mathematical knowledge from a set of physical materials by the actions he performs on or with the materials.

*Sample items are given in the Appendix.

2. Mathematical concepts should not be presented to young children through the use only of the symbols of mathematics or verbalizations. Explanations which accompany the child's actions, however, may facilitate his acquisition of mathematical concepts.

3. There should be a continuous interplay between the spoken words which symbolize a mathematical concept and a set of actions a child performs while constructing something that makes the concept tangible.

4. In order to teach a concept, it is necessary to use different assortments of physical materials and different types of activities all of which are related to the development, by the child, of the same concept(s).

5. The principle of reversibility should be employed (i.e., returning a transformed set of conditions to an original set of conditions).

6. Situations must be contrived in which the children are led to multiple focusing (e.g., if A is the same length as B, then B is also the same length as A).

7. Situations must be contrived which involve more than one child so that the children may interact.

8. The principle of equilibration should be employed.

Measuring instruments. Five instruments were constructed to measure pupil capabilities. The first instrument, the Length Comparison Test, was designed to measure the ability of children to establish a length relation between two curves. Six different material sets were used. Three items, one a "longer than," one a "shorter than," and one a "same length as" item, were presented to the child in the case of each material set for a total of 18 items.

The second instrument, the Conservation of Length Relations Test, consisted of two parts. In the first part of each of the 18 items, the children were asked to compare the lengths of two curves. Since the material used in the items differed from those in either Instructional Sequence I or the Length Comparison Test, these 18 first parts were considered as an Application Test for Instructional Sequence I (hereafter called the Length Comparison Application Test). The second part of each item involved the ability of the child to conserve the length relation he had just established. Nine of the items also involved the ability of the child to use the asymmetric property of "longer than" and "shorter than" or logical consequences. These nine items comprised an instrument which will be designated as the Conservation of Length Relations: Level II Test. The remaining nine items comprise an instrument which will be designated as the Conservation of Length Relations: Level I Test.

"Yes" was the correct response in the case of the nine items of the Level I Test. "No" was the correct response for each of the nine items of the Level II Test. The children were required to respond in the presence of a perceptual conflict so that if a child based his response on visual perception he would give an incorrect response. Three distinct length-preserving transformations were used to produce the perceptual conflicts. Also, different material sets were utilized.

The third instrument, the Reflexive and Nonreflexive Test, consisted of six items of a diversified nature. Three of the items involved the reflexive property of "the same length as" and three items involved the nonreflexive property of "longer than" or "shorter than." Five different material sets were employed. "Yes" was the correct response to the items involving the reflexive property, and "No" was the correct response to the remaining three items.

The fourth instrument, the Transitivity Test, consisted of six items where "Yes" was the correct response for three items. For these items, each of the relations "longer than," "shorter than," and "the same length as" was included. "No" was the correct response for the remaining three items. Each of the latter three items involved transitivity of "the same length as." It was not possible for the child to use a non-transitive hypothesis to arrive at a correct response because all of the perceptual cues were biased against a correct response and the child was not allowed to directly compare the two curves under consideration.

Instructional and Evaluational Sequence

Small group instructional procedures were utilized in each room. An instructional group generally consisted of six children with teacher aides present to guide the remaining children. After the Length Comparison Test was administered, Instructional Sequence I was administered for a sequence of seven sessions of 20-30 minutes per session. Due to small-group instructional procedures, the total instructional time spanned more than seven days for any one class. However, any one child was involved in only seven instructional sessions. The Length Comparison Test, the Length Comparison Application Test, the Conservation of Length Relations Test, the Reflexive and Nonreflexive Test, and the Transitivity Test were administered during the days immediately following the last instructional session. The Length Comparison Test was not administered a second time to one class because that class earned a high mean score on the first administration of the test. The material in Instructional Sequence I was administered to that class, however, to support the interpretation of the remaining tests.

Instructional Sequence II and III began immediately after the testing period following Instructional Sequence I. Instructional Sequence II was administered in three sessions of 20-30 minutes per session and Instructional Sequence III was administered in five sessions of 20-30 minutes per session. The second administration of the Length Comparison Application Test, the Conservation of Length Relations Test, the Reflexive and Nonreflexive Test, and the Transitive Test began one day after the last instructional sequence.

Testing Procedures

The children were tested on a one-to-one basis. The items were assigned at random by test to each child so that each had a different sequence of the same items. All tests were administered by specially trained evaluators.

The Length Comparison Test was scored on a basis of the number of correct comparisons a child was able to perform. The Conservation of Length Relations Test was administered at one sitting so that a child would be forced to respond "Yes" or "No" in a random sequence. If a child established a relation, regardless of whether he established a "correct" or "incorrect" relation, he was tested on his ability to conserve that relation. In the case of the Transitivity Test, unless a child established two correct comparisons, no measure was obtained on his ability to use the transitive property of that relation.

Design

Length Comparison Test. An analysis of variance technique was used to study the profiles of four- and five-year-old children with regard to the Length Comparison Test on the first and second administration (Greenhouse and Geisser, 1959). In particular, the design allows for testing of the hypothesis that the profile of mean scores on the first and second administration does not differ for the four- and five-year-olds as well as providing a test for possible differences in the mean scores on the first and second administration across age. An item analysis was also conducted.

Conservation of Length Relations Test. The Conservation of Length Relations Test consisted of nine items for which a response of "Yes" was correct and nine items for which a response of "No" was correct. One may think of each student's response set as being an ordered 18-tuple where each element is either "Yes" or "No." If each response set is considered to be a random sample from 2^{18} such response sets, it has probability of 2^{-18} of occurring (Feller, 1957, p. 29). If a child guessed during the test, then one may consider his responses as being nothing more than an 18-tuple of "Yes's" or "No's" for elements, where "Yes" or "No" for any one entry each had probability of 1/2 of occurring. In this case, his response set may be considered as a random sample, and the probability he obtained at least six correct "Yes" responses and six correct "No" responses is not greater than .06.

For a child to be classified as being able to conserve length relations and conserve length relations involving the asymmetric property and logical consequences, he then must have at least six of the nine items which were written to exemplify Level I and six of the nine items which were written to exemplify Level II correct. In such case, the child is said to meet criterion for Level I and II.

If one considers the nine items written at either Level I or Level II regardless of the nine items written at the other level, a probability of only approximately .02 exists that a child responded correctly to eight or nine items at that Level if he guessed. Thus, if a child does not meet criterion for Level I and II one may consider his responses to one of the two items sets written at Level I or Level II. Clearly, a high probability exists that those children who scored at least eight or nine correct for a particular item set may have responded to those items on a basis other than guessing. Such children may be candidates for being classified at just Level I or Level II. One cannot, however, with any degree of confidence, assert that in fact such children did not possess a response bias unless the remaining nine items are considered. For example, if a child was able to score an eight or nine on Level I items and responded on a basis of guessing on Level II items, then a probability of only .02 occurs that the child had at most one correct "No" response. If this unlikely event occurred, whether a response bias existed or whether the child responded on the basis of the perceptual cues is an open question. For a child to meet criterion for just Level I or Level II then, he must respond correctly to eight or nine items of the Level in question and no less than two and no more than five items of the other level. It must be pointed out that the criterion is a conservation one since it is known that children do respond on the basis of perceptual cues (Steffe, 1966).

A principal component analysis was conducted using all 18 items to aid in the interpretation of the above criteria. Item difficulties were determined as well as internal consistency reliabilities, both of which also can be used in interpretation of the criteria established. In order to check the hypothesis that the distribution of total scores did not differ from a theoretical distribution based on random responses, a "goodness of fit" test was employed (Seigel, 1956, pp. 42-46).

In order to detect any significant changes in the number of children meeting criterion on the conservation of length relations test for Level I and II from the first to second administration, the McNemar test for significance of changes was used. (Seigel, 1956, pp. 63-67). Thus, those children meeting criterion were given a "1" and those not meeting criterion were given a "0," so only nominal scale of measurement was employed. According to Seigel, the "McNemar test for the significance of changes is particularly applicable to those 'before and after' designs in which each person is used as his own control in which measurement is in the strength of either a nominal or ordinal scale (p. 63)." Explicitly, the null hypothesis is: H_0: For those children who change, the probability P_1 that any child will change from C (criterion) to ~C (noncriterion) is equal to the probability P_2 that he will change from ~C to C. The alternative hypothesis is: H_1: $P_1 < P_2$.

Reflexive and Nonreflexive Test. The set {Yes, No} represents possible responses for the six items of the Reflexive and Nonreflexive Test. Other responses were possible, but they occurred with zero probability in the testing sessions. Because there are 2^6 different sextuples with "Yes" or "No" as elements, if a child guessed, the probability that any one of the 2^6-tuples occurred is 2^{-6}. Under these conditions, the probability of a child obtaining at least five or six correct responses is approximately .11. It must be pointed out, however, that children do respond on the basis of perceptual cues, so that the actual probability that a child who does not possess the ability to conserve length could obtain five or six may be much lower than .11.

If a child responded on the basis of a bias (always says "Yes" or "No"), then he would not obtain a five or six. Moreover, if a child possesses only the ability to use either the reflexive or nonreflexive property, he also would not achieve a five or six. Hence, the performance criterion of a total score of five or six was established. A "goodness of fit" test was employed to test the hypothesis that the distribution of total scores does not differ from a distribution based on random responses. In order to detect any significant changes in the number of children meeting criterion from the first to second administration, the McNemar test for significance of changes was used. Explicitly, the null hypothesis is H_0: For those children who change, the probability P_1 that any child will change from C to ~C is equal to the probability P_2 that he will change from ~C to C. The alternative hypothesis is: H_1: $P_1 < P_2$.

Transitivity Test. Based on the average item difficulty of the Length Relations Application Test, a parameter was established (average item difficulty) which may be regarded as an efficiency level of the child's ability to establish length relations between curves. Using this parameter, r, the probability that a child could establish a correct relation in each of the two necessary overt comparisons on any item (comparisons between A and B and between B and C, where ARB and BRC and R is the relation) in the Transitivity Test was r^2. The calculation assumes that the comparisons are performed independently.

If a child responded on a random basis to a relational question concerning A and C (such as, "Is A longer than C?"), the probability p of a correct response on any item is $r^2/2$. Using this value of p, a performance criterion was established and a "goodness of fit" test performed on the distribution of total scores to the theoretical distribution of total scores based on guessing.

To establish whether the ability to conserve length relations is necessary (sufficient) to enable children to use the transitive property, an inspection was made of those children who met criterion on each test instrument. If the ability to conserve length relations is necessary for the ability to use the transitive property, then each child who attains criterion on the Transitivity Test must also meet criterion on the Conservation of Length Relations Test. If the ability to conserve length relations is sufficient for the ability to use the transitive property, then each child who meets criterion on the Conservation of Length Relations Test must also meet criterion on the Transitivity Test. Other interdependencies were investigated in the same way.

Results of the Study

The results of the study are partitioned as follows: Length Comparison Test; Length Comparison Application Test; Conservation of Length Relations Test; Reflexive and Nonreflexive Test; Transitivity Test; and Conservation and Transitivity Relationships.

Length Comparison Test

The results of an internal-consistency reliability study (Table 3) revealed that the reliabilities associated with the total test scores were quite substantial and support analyses of the data. In the case of the first administration, the reliabilities for the subsets were also substantial. For the second administration, however, the reliability for the six items which were designed to measure the ability of children to establish the relation "shorter than" was low. Various reasons may be given, the most apparent of which is the high mean and relatively small standard deviation (Table 4). It is known that easy tests may be unreliable.

No differences were statistically discernible for the variable Age using the total scores as the dependent measure (Table 5) although the mean score for the second administration was significantly greater than the mean score for the first administration. No interaction of Age and

Table 3
Reliabilities of Length Comparison Test: First and Second Administration
(Kuder-Richardson #20)

Test	Reliability	
	First Administration	Second Administration
Total	.91	.83
Longer Than	.82	.71
Shorter Than	.87	.43
Same Length As	.77	.73

Table 4
Mean and Standard Deviations of Length
Comparison Test: First and Second Administration

Test	First Administration		Second Administration	
	$\overline{X}$	S.D.	$\overline{X}$	S.D.
Total	10.68	5.35	14.55	3.53
Longer Than	4.38	1.91	4.94	1.43
Shorter Than	3.29	2.32	5.12	1.07
Same Length As	3.00	2.02	4.49	1.70

Table 5
ANOVA Summary
Length Comparison Application Test

Source of Variation by Test	F
Total Scores	
A (Age)	2.65
B (Tests: First vs. Second Administration)	22.45**
AB	< 1
Longer Than	
A (Age)	< 1
B (Tests)	2.85
AB	< 1
Shorter Than	
A (Age)	3.42
B (Tests)	26.35**
AB	< 1
Same Length As	
A (Age)	2.04
B (Tests)	14.18**
AB	2.80

**$p < .01$

Tests occurred that indicate that the difference between the means for each group on the second administration was not significantly different than the differences between the means for each group on the first administration.

On the subtest "longer than," the children started with relatively high mean scores (68 and 75 percent for the four- and five-year-olds, respectively) and ended with mean scores 78 and 86 percent, a nonsignificant gain, statistically. In the case of the subtest "shorter than," a large gain was noted for both the four- and five-year-olds (from 43 to 76 percent for age four and from 62 to 91 percent for age five). Again, Age was not significant. In the case of the subtest "same length as," a substantial increase was again present (48 to 57 percent for age four and 46 to 80 percent for age five). Age was again nonsignificant as was the interaction of Age and Tests. On the basis of the test scores alone, one may hypothesize that an interaction occurred. The nonsignificant interaction may be due to the power of the statistical test involved.

All the correlations computed between test scores on the first and second administration with the variables Verbal Maturity, I.Q., Age, and Social Class were low, although some differed significantly from a zero correlation. Age correlated significantly (Total test--.42, Shorter Than--.41, Same Length As--.42; $p < .02$) with scores on the second administration except for the subtest "longer than." The correlation coefficient between Social Class and the subtest "same length as" on the first administration was also statistially significant (-.41, $p < .02$) but negative.

Length Comparison Application Test

In order to ascertain whether the ability of children to compare lengths of curves was restricted to six specific material sets, the Length Comparison Application Test was administered and a correlation study conducted using scores of eight tests (the total tests and subtests thereof for the second administration of Length Comparison Test and for the first administration of Length Comparison Application Test). The correlation of .81 between total scores (Table 6) along with the significant pair-wise correlation of the respective subtests indicates a high degree of relationship.

Table 6
Correlation Matrix
Length Comparison Test (Second Administration) and
Length Comparison Application Test (First Administration)

Test	1	2	3	4	5	6	7	8
LCT								
1. Total	1.00	.82**	.79**	.88**	.81**	.71**	.78**	.65**
2. Longer		1.00	.52**	.53**	.77**	.71**	.77**	.58**
3. Shorter			1.00	.57**	.69**	.60**	.65**	.59**
4. Same As				1.00	.58**	.50**	.56**	.48**
LCAT								
5. Total					1.00	.83**	.93**	.89**
6. Longer						1.00	.79**	.60**
7. Shorter							1.00	.72**
8. Same As								1.00

**Significantly different from zero correlations; $p < .01$

The Length Comparison Application Test was administered twice, once before the completion of Instructional Sequences II and III and once after. The pupil performances on the second administration are of interest because Instructional Sequence III contained additional exercises on length comparisons. However, no apparent changes in the mean scores were observable across administrations (Table 7). All the reliabilities on the first administration were substantial (Table 8). On the second administration, the

Table 7
Means and Standard Deviations of Length Comparison Application Tests

Test	First Administration $\bar{X}$	First Administration S.D.	Second Administration $\bar{X}$	Second Administration S.D.
Total	14.10	3.89	14.44	3.14
Longer Than	5.02	1.39	5.16	1.22
Shorter Than	4.34	1.81	4.86	1.03
Same Length As	4.74	1.50	4.42	1.58

reliability of the subtest "shorter than" was very low. A high mean score and small standard deviation may contribute to this reliability.

Table 8
Reliabilities of Length Comparison Application Tests
(Kuder-Richardson #20)

Test	Reliability: First Administration	Reliability: Second Administration
Total	.85	.76
Longer Than	.71	.63
Shorter Than	.77	.18
Same Length As	.68	.65

The correlation of the variables Verbal Maturity, I.Q., Age, and Social Class with the total test scores and subtest thereof on the first and second administration were low (-.12 to .39). All but one of the significant correlations involved Age. This is consistent with the correlations reported earlier for the Length Comparison Test.

Conservation of Length Relations Test

An internal consistency reliability study was conducted for each test administration. The range of the reliabilities was .81 to .88. The item difficulties (Table 9) and means (Table 10) for the Conservation of Length Relations Tests indicated that performance of children on Level I and Level II items was similar for the first test administration. There was a major difference, however, for the second administration in that Level I items were considerably less difficult that Level II items, items which remained at about the same difficulty level for both test administrations.

A principal component analysis was conducted (Table 11). Factor 1 of the first test administration was a bipolar factor where the items at Level I loaded negatively and the items at Level II loaded positively. Four of the five positive loadings which exceeded .5 were items involving the asymmetrical property of "longer than" or "shorter than." The remaining

Table 9
Item Difficulty of Conservation of Length Relations Tests

	Item	Difficulty: First Administration	Difficulty: Second Administration
Level I	1	.59	.83
	2	.39	.69
	3	.55	.85
	4	.49	.87
	5	.51	.83
	6	.51	.77
	7	.37	.73
	8	.49	.85
	9	.39	.73
Level II	1	.47	.46
	2	.43	.58
	3	.49	.44
	4	.59	.58
	5	.43	.44
	6	.43	.52
	7	.57	.56
	8	.43	.46
	9	.49	.58

Table 10
Means and Standard Deviations of Conservation of Length Relations Test: Level I and Level II

Level	First Administration $\bar{X}$	First Administration S.D.	Second Administration $\bar{X}$	Second Administration S.D.
Level I	4.29	3.17	7.13	2.56
Level II	4.33	2.80	4.62	2.92

positive loading which exceeded .5 involved the statement, "If A is the same length as B, then A is not longer than B." The six items which had loadings greater than .5 for Factor 2 of the first test administration included four items which involved logical consequences of the relations, one of which involved the asymmetrical property of "shorter than" and one of which tested conservation of "the same length as."

Of the two identifiable factors of the second test administration, the items which had loadings greater than .5 were all Level II items for Factor I and Level I items for Factor 2. Moreover, each Level II item had a loading greater than .5 on Factor 1. These factors clearly may be a result of the item difficulties for the second test administration.

Table 11
Principal Component Analysis of Conservation of Length Relations Tests

Item Level		First Administration 1	First Administration 2	Second Administration 1	Second Administration 2
Level I					
	1	-.5179	.4928	-.1938	.6206
Longer Than	2	-.8139	.3029	.0778	.7065
	3	-.7534	.1310	.3150	.4925
	4	-.6982	.3444	-.2794	.5122
Shorter Than	5	-.6776	.4312	-.2651	.4185
	6	-.7831	-.0674	-.1996	.7985
	7	-.6903	.5853	-.3807	.8069
Same Length As	8	-.7261	.2393	-.3466	.3631
	9	-.8440	.4050	-.4118	.7599
Level II					
	1	.5748	.4854	.5134	.3771
Longer Than	2	.3730	.5885	.7704	.2654
	3	.6523	.0404	.8291	.2600
	4	.8592	.2577	.5999	.1304
Shorter Than	5	.7244	.5162	.7887	.2206
	6	.4528	.5671	.5580	.0838
	7	.5512	.6195	.7462	.2527
Same Length As	8	.3729	.2845	.6475	.2374
	9	.3865	.7055	.7429	.1681
Percent Communality		45.57	21.61	41.67	35.00

These principal component analyses justified identifying two levels of items in the Conservation of Length Relations Test.

Level I criterion was met by one four-year-old and three five-year-olds for the first test administration. Of these four children, only one five-year-old met criterion for Level I and II on the second test administration, and the four-year-old met criterion for Level I. The remaining two children did not meet any criterion on the second test administration. A total of four four-year-olds and six five-year-olds met criterion for Level I on the second test administration.

In case of the first test administration, seven children scored eight or nine on Level II items but at most five on Level I items. Of these seven, five had only a zero or one on Level I items. One of the remaining two children met criterion for Level I and II on the second test administration. The other child did not meet any criterion on the second test administration. No child met criterion for Level II only on the second administration.

After Instructional Sequence I, three four-year-olds and three five-year-olds met the criterion for Level I and II. Five four-year-olds and fourteen five-year-olds met the same criterion after Instructional Sequence II and III. According to the McNemar Test, the change for the five-year-olds was significant ($\chi^2 = 6.67$; $p < .01$), but the change for the four-year-olds was not.

If children do, in fact, respond "Yes" or "No" in a random fashion to items presented, then the distribution of total scores should not depart from a binomial frequency distribution based on random responses, except for chance fluctuations. The actual frequency distribution of scores for each group by Conservation Level did statistically depart from a theoretical distribution at the .01 level (Table 12).

Table 12
Comparison of Theoretical and Actual Frequency Distribution
Conservation of Length Relations

Test By Age	χ^2
Four-Year-Olds	
Level I, First Administration	744.5**
Level I, Second Administration	1712.8**
Level II, First Administration	140.8**
Level II, Second Administration	69.6**
Five-Year-Olds	
Level I, First Administration	351.3**
Level I, Second Administration	2071.5**
Level II, First Administration	267.0**
Level II, Second Administration	495.9**

**p < .01

The correlation between the Level I and Level II total scores with Verbal Maturity, Age, and Social Class was not significantly different from a zero correlation except for the correlation between I.Q. and the Level II total scores on the second test administration. This correlation of .34 was low.

Reflexive and Nonreflexive Test

The reliabilities for the Reflexive and Nonreflexive Test were .43 for the first test administration and .53 for the second test administration. A contribution to the low test reliabilities was the existence of more than one factor in the test (Table 13). In case of each test administration, the items loaded on two factors. Factor 1, in case of the first test administration, was a combination of the reflexive property and the type of transformation and Factor 2 was a combination of conservation involving the nonreflexive property and the type of transformation. It is noted that for both factors, two of the items involving the same property loaded with a higher value than the third. The third item always involved a different transformation from the two others, which involved the same transformation. These factors clearly justify the criterion established because for a child to know that a curve is the same length as itself, he must also know it is not longer or shorter than itself.

Table 13
Principal Component Analysis for the Reflexive and Nonreflexive Test

Item	*First Administration* 1	*First Administration* 2	*Second Administration* 1	*Second Administration* 2
Nonreflexive				
1	.1971	.6889	.7808	.0553
2	-.1639	.6452	.8748	.0444
3	.1975	.3252	.5354	-.0007
Reflexive				
4	-.9071	-.0168	-.0819	.7808
5	-.9341	.1090	-.0586	.8334
6	-.4188	.0184	.0940	.3311
Percent Communality	52.63	26.88	42.45	35.84

The item difficulties for the first test administration ranged from .24 to .51 with four item difficulties below .40. The item difficulties for the second test administration ranged from .37 to .88 with only one difficulty below .40. All of the item difficulties increased from the first to second administration with the greatest increase for the items involving the reflexive property. A change from 2.12 to 3.75 in the means for the first and second administration reflected the modest increase in item difficulty.

One four- and one five-year-old earned a score of five or six on the first test administration. The number of four- and five-year-olds meeting the criterion on the second test administration increased to six and nine, respectively. The two children who met the criterion on the first test administration did not meet the required level of performance on the second administration. However, the number of students who changed from noncriterion to criterion was greater than the number of children who changed from criterion to noncriterion ($\chi^2 = 8.50$, $p < .01$). There was also an increase in the number of children that responded correctly to all the reflexive items but did not meet the criterion. The change was from seven to twenty-one from the first to the second test administration.

When the distribution of total scores by the four-year-olds was considered, it was found that the frequency distribution for both the first ($\chi^2 = 32.8$) and second ($\chi^2 = 50.7$) test administration departed statistically at the .01 level from a binomial distribution. The theoretical and actual frequency distributions of scores earned by the five-year-olds on the first ($\chi^2 = 37.7$) and second ($\chi^2 = 31.3$) test administration also departed statistically at the .01 level.

All correlations of the total scores with the variables Verbal Maturity, I.Q., Age, and Social Class were low. However, the correlations between total scores and Social Class were significantly different from zero.

Transitivity

The reliabilities for both test administrations for the Transitivity Test (.50 and .45) were low. This may be expected because the principal component analysis (Table 14) revealed the existence of more than one factor in the test. For the first test administration, two items (one involving transitivity of "shorter than" and one "longer than") loaded greater than .5 on Factor 1. Factor 2, first administration, is a combination of transitivity of "longer than" and "same length as." For the second test administration, Factor 1 involved transitivity of "same length as," and Factor 2 involved transitivity of "shorter than" and "longer than," a clear dichotomy.

Table 14
Principal Component Analysis for the Transitivity Test

Item	Relation	First Administration 1	First Administration 2	Second Administration 1	Second Administration 2
1	Same Length	.4362	.4998	.8001	-.0065
2	Same Length	.2940	.4637	.5762	.3157
3	Shorter than	.6970	-.1967	.1468	-.4648
4	Longer than	.5621	-.5222	.1204	-.5562
5	Same Length	.3424	.2903	.4260	-.0658
6	Same Length	.4241	-.0545	.7806	-.1921
Percent Communality		44.44	27.90	61.50	11.11

The mean scores for the first and second administration were 2.00 and 2.67, respectively, an increase which reflected the increase in item difficulty for each item. Only one item on any administration had a difficulty that exceeded .5. Four four-year-olds and five five-year-olds met the criterion (a total score of five or six) for transitivity on the first test administration. A total of fifteen students met the criterion on the second administration of which five were four-year-olds and ten were five-year-olds. Five students that met the criterion on the first test administration did not meet the criterion on the second administration. Three of these students were unable to make the necessary length comparisons upon which to base the transitive property. Therefore, only two students may have lost transitivity. The level of performance of one of these two students may involve a chance fluctuation since transitivity was exhibited three out of five times on the second test administration.

The actual frequency distributions of scores earned by the four-year-olds on the first ($\chi^2 = 6.5$) and on the second ($\chi^2 = 9.1$) test administrations did not depart statistically at the .05 level from a binomial distribution based on random responses. Since the actual frequency distributions for the four-year-olds did not depart from the theoretical distribution, no four-year-olds were considered to have the ability to use the transitive property of the length relations involved. In the calculation of the theoretical binomial distribution based on guesses, a probability p for correct responses was .30. This value is based on an efficiency level of .78 as calculated from the Length Comparison Application Test, first test administration.

The actual frequency distribution of scores earned by the five-year-olds on the first ($\chi^2 = 17.2$) and second test administration ($\chi^2 = 74.9$) did depart statistically from a binomial distribution at the .01 level. The main departure for the first administration scores was in the number of 0 and 1 scores. An increase in frequency of total scores in the range of 3 to 6 was noted in the results of the second test administration.

An investigation of characteristics of children not meeting the criterion and those meeting the criterion revealed little difference between the mean age for the two levels of performance of any one group. There was a small difference between the mean Verbal Maturity scores and for the mean I.Q. scores. Correlation between the variables Verbal Maturity, I.Q., Age, and Social Class and the levels of performance were not statistically significant.

Relationships Among the Variables

On the first test administration, two children met criterion on the Reflexive and Nonreflexive Test. Neither of these children met criterion on the Transitivity Test or criterion on Conservation of Length Relations: Level I and II Test. One child met criterion only in the case of the Conservation of Length Relations: Level I Test. On the second test administration, only one child out of the 14 who met criterion on the Reflexive and Nonreflexive Test met criterion on the Transitivity Test. This child did not meet the criterion on the Conservation of Length Relations: Level I Test or on the Conservation of Length Relations: Level I and II Test. However, seven of the 14 children who met criterion on the Reflexive and Nonreflexive Test met criterion on the Conservation of Length Relations: Level I and II Test, and three children met criterion in the case of Conservation of Length Relations: Level I Test.

Four children met criterion of Length Relations: Level I Test but not on the Level II Test for the first test administration. Two of these children met criterion on the Transitivity Test and one met criterion on the Reflexive and Nonreflexive Test. The two children meeting criterion on the Transitivity Test did not meet criterion on the Reflexive and Nonreflexive Test. Only four out of the ten children who met criterion on the Conservation of Length Relations: Level I Test but not the Level II Test on the second test administration met criterion on the Reflexive and Nonreflexive Test. None of the ten met criterion on the Transitivity Test.

On the first test administration, only one out of the six children who met criterion on the Conservation of Length Relations: Level I and II Test met criterion on the Transitivity Test. None of these six children met the criterion on the Reflexive and Nonreflexive Test. On the second test administration, seven of the 19 children who met criterion on the Conservation of Length Relations: Level I and II Test met criterion on the Reflexive and Nonreflexive Test. Seven different children met criterion on the Transitivity Test. Five children did not meet criterion on the Reflexive and Nonreflexive Test or on the Transitivity Test.

On the first test administration, only two of five children who met criterion on the Transitivity Test met criterion on the Conservation of Length Relations: Level I Test. One of these five children met criterion on the Reflexive and Nonreflexive Test. On the second test administration, seven of the ten students who met criterion on the Transitivity Test also met criterion on the Conservation of Length Relations: Level I and II Test, but only one child met criterion on the Reflexive and Nonreflexive Test.

Conclusions, Discussions, and Implications

Length Comparison Test

Before or after Instructional Sequence I on length comparison, the performance of four-year-old children in establishing a relation between two curves is not different from that of five-year-olds. It appears that four- and five-year-old children easily learn the relation "longer than" through informal experiences or testing facilitates learning of this relation. Beilen and Franklin (1962) did find that testing facilitates first-grade children's acquisition of measurement tasks which, in addition to the fact that testing was conducted for all three relations, leads to the conclusion that children acquire "longer than" through informal experience to a greater extent than "shorter than" or "the same length as."

Instruction on establishing length comparisons does significantly improve the ability of both four- and five-year-old children to establish length comparisons. The instructional experiences utilized in this study involved a continuous interplay between language and manipulation of objects as Bruner and Kenney (1964) recommend. This interplay was an endeavor to eliminate experiences dependent solely upon language and not real practical action, which Adler (1964) considers a failure of formal education and which should have aided the children in not responding on a perceptual basis when establishing length relations as suggested by Wohlwill (1960).

The ability of four- and five-year-old children to make length comparisons involving the relations "longer than," "shorter than," and "same length as" is not limited to situations in which they learned to establish these relations -- as the children had the ability to use the relations in novel length comparison situations. The formal experiences with concrete materials was sufficient for a majority of the children to reach an overt operational level with length comparisons, a level of performance which was retained over the several months this study was in progress.

There appears to be little, if any, relation between the variables of Verbal Maturity, I.Q., Age, and Social Class and the ability of four- and five-year-old children to make length comparisons involving "longer than," "shorter than," and "same length as." This is similar to Beilen and Franklin's (1962) findings that I.Q. was not a factor in first-grade children's learning to measure length.

Conservation of Length Relations

The definitions given for length relations and conservation of these relations (i.e., that the relation obtains regardless of the proximity of the curves) seem to have been supported by the results of the study. On the Length Comparison Application Test, first administration, the mean score was 78 percent (14.10 out of 18 items) with a standard deviation of only 3.89. At this point in time, the children in the study were able to associate a relational term with an overt comparison of curves in such a way that they were able to discriminate among the comparisons denoted by "longer than," "shorter than," and "the same length as." The particular relation a child established on the first administration of the applicati n test through overt comparison was a function of the proximity of the urves involved. This is supported by the fact that at most two children could be classified at only Level I and at most four children could be classified as meeting criterion on the Conservation of Length Relations: Level I and II Test (those four who met criterion on both the first and second administration of the Conservation of Length Relations Test). With the exception of these last four children and possibly the former two, there is no evidence that at the time of the first administration of the Length Comparison Application Test an overt comparison constituted a logical-mathematical experience for the child making the comparison. The overt comparison was certainly not sufficient for the child (using Piaget's terms) to disengage the structure of the relation he established. It certainly may be the case that the relation for the child not only was a function of the proximity of the curves but was a function of the external physical situation so that the child did not think about the relation in the absence of the external situation. In Bruner and Kenny's (1964) terms, the child had not internalized the relation; or in Lovell's (1966a) terms, the child was not aware of the significance of his actions in the overt comparison of the curves.

The definitions of Level I and Level II were well supported by the principal components analysis on the first administration. This analysis shows that the items written at Level I and Level II involve differential abilities. In particular, for the pretest, the items written at Level II which involved the asymmetrical property of "longer than" or "shorter than" loaded on Factor 2 as well as an item involving a logical consequence of "the same length as." On the second test administration, the items written at Level I were much less difficult than those written at Level II, which certainly contributed to the factors present in the principal component analysis.

Level I items were constructed to measure the extent to which the children realize that the length relation they established between two curves is independent of the proximity of the curves. As noted, before the administration of Instructional Sequences II and III, only about 12 percent of the children could be categorized at Level I. After the administration of Instructional Sequences II and III, however, the evidence indicated that about 57 percent of the children could be categorized at that Level. At the same two points in time, the percents were 8 and 37 with regard to Level I and II, which was a statistically significant change. It must be emphasized that the children in this 37 percent not only were able to establish a relation between two curves and retain the relation regardless of the proximity of the curves but were able to use the asymmetric property and logical consequences of the relation under consideration. It is certainly true that the experiences contained in Instructional Sequences II and III did not readily increase the children's ability to use logical consequences of the relations they were able to establish.

The data suggest that the mean I.Q. for the five-year-old children who met criterion for Level I and II was greater than the mean I.Q. for those who did not meet criterion. The correlation of total scores for Level I and II with the variables of Verbal Maturity, I.Q., Age, and Social Class were not significant with the possible exception of a low correlation between I.Q. and Level II posttest scores.

Reflexive and Nonreflexive Properties

Very few four- and five-year-old children were able to use the reflexive and nonreflexive properties on the first administration of the Reflexive and Nonreflexive Test. Elkind (1967) apparently would identify the ability to use the reflexive property as conservation of identity even though he did not subdivide conservation of identity with regard to the reflexive and nonreflexive properties. An effort is made here not to confuse conservation of identity with the ability to use the reflexive and nonreflexive properties nor to confuse conservation of length with conservation of length relations.

Some four- and five-year-old children have the ability to use the reflexive property but not the nonreflexive property. Instructional experience on length comparisons appear to be sufficient for such children to exhibit the reflexive property, as 14 percent of the sample were able to use the reflexive property on the first test administration as compared to four percent who were able to use both properties.

Instructional Sequences II and III significantly increased the ability of four- and five-year-old children to use both properties. On the second test administration 41 percent of the sample were able to use only the reflexive property and 30 percent of the sample were able to use both. Only 29 percent of the sample did not display an ability to use the reflexive or nonreflexive properties. These conclusions substantiate Piaget's theory that experience is a necessary but not a sufficient condition for the development of logical thought processes because all the children received the same selected experiences. Certainly the data substantiate that the ability to use the reflexive property is different from and precedes the ability to use the nonreflexive property.

There appears to be little, if any, relation between the student variables Verbal Maturity, I.Q., Age, and Social Class and scores earned by four- and five-year-old children on the Reflexive and Nonreflexive Test. Only correlations involving Social Class were significantly different from zero, but these correlations were low.

Transitivity

Few five-year-old children were able to use the transitive property after only instructional experience in establishing length relations. At this point in time, only 16 percent of the five-year-olds used the transitive property. At the same point in time, the distribution of total scores for the four-year-olds did not statistically depart from a binomial distribution based on random responses, so no four-year-old was considered able to use the transitive property of length relations. Some children performed poorly because of their inability to establish the two initial comparisons, an inability Smedslund (1963b) considers as a reason for failure of some young children to use the transitive property.

Instructional Sequences II and III did increase the ability of five-year-olds to use the transitive property, since the percent of five-year-olds able to use the transitive property increased to 31. These same experiences did not increase the ability of four-year-old children to use the transitive property because again the distribution of total scores for the four-year-olds did not statistically depart from a binomial distribution based on guessing. The number of five-year-olds that used transitivity of length relations is below that found by Braine (1959) but above that found by Smedslund (1964). It appears that these experiences were not logical-mathematical experiences that readily increase children's ability to use the transitive property. All the children may not have had a mental structure sufficient to allow assimilation of the information.

The mean Verbal Maturity and I.Q. of five-year-old children who were able to use the transitive property appeared to be slightly higher than for those who do not use this property. However, the correlations between these two variables and transitivity scores earned by the total sample was not statistically different from zero. Also, there appears to be little, if any, relationship between the variables Age and Social Class and the ability of four- and five-year-old children to use the transitive property.

Relationships Among the Variables

The relationships of reflexive and nonreflexive properties, conservation of length relations, and transitivity of length relations will be discussed on each of the first and second test administrations.

On the first test administration only two children met criterion on the Reflexive and Nonreflexive Test so that a discussion of relationships is not appropriate. However, on the second administration, 30 percent of the children met criterion. Of this 30 percent, only one child met criterion on the Transitivity Test. Because there were 10 children who met criterion on the Transitivity Test, it is quite apparent that the ability to use the reflexive and nonreflexive properties as measured here is not a necessary or a sufficient condition for the ability to use transitivity of length relations. This observation is quite consistent with the fact that the reflexive property of "the same length as" does not imply the transitive property of "the same length as" nor does the nonreflexive property of "longer than" or "shorter than" imply the transitive property of these two relations, on a logical basis. Conversely, the transitive property of "longer than" or "shorter than" does not imply the nonreflexive property of these two relations. Because on a logical basis the reflexive property of "the same length as" is (under restricted conditions) a consequence of the symmetric and transitive properties of "the same length as," and because some children could use the reflexive property but not the transitive property, there may be factors which enable children to use the reflexive property before they are able to use transitivity (e.g., spatial imagery or the definition of "the same length as"). In fact, the results indicate that the reflexive property may be necessary for transitivity. This observation may be due to the possibility that use of the reflexive property in this study was more of a "learned response" than a logical-mathematical process.

It also appears that use of the reflexive and nonreflexive properties is not a necessary or sufficient condition for being able to conserve length relations. Of the 30 percent who met criterion on the Reflexive and Nonreflexive Test, only seven children met criterion for the Conservation of Length Relations: Level I and II Test. This observation is consistent with the logical relationships of the properties of the relations. However, the data do not contradict the fact that being able to use only the reflexive property may precede an ability to conserve length relations at Level I and, therefore, Level II. The data of this study support the contention that conservation of identity is not unitary in nature. Certainly, if a child judges that a stick is the same length as itself, he must also judge that it is not longer or shorter than itself, or a contradiction would be present. On a logical basis and on a psychological basis, when one considers "conservation" problems, it is necessary to consider the properties of the relations which may be involved.

For those nineteen children who met criterion on the Conservation of Length Relations: Level I and II Test, seven met criterion on the Transitivity Test. Since only ten children met criterion on the Transitivity Test, it seems that conservation of length relations: Level I and II is necessary for transitivity. The fact that two of three children who met criterion on the Transitivity Test but not for conservation length relations: Level I and II, did not meet criterion for conservation of length relations: Level I or for reflexivity and nonreflexivity, indicates an inaccurate assessment. The above data are consistent with Smedslund's (1963b) observation that what he calls conservation of length is a necessary condition for what he calls transitivity.

The study involves implications for further research and development. Among these implications, the following are relevant. (1) With the exception of the transitive property, it may be important to first introduce the properties, relationships, and consequences of the relations involved at the point in time in which the children are first able to associate a relational term with an overt comparison and before perceptual conflict is introduced. The children could then observe, with perceptual support, the properties, etc., involved. If the children were thus able to learn that the relation(s) they establish is (are) not a function of the proximity of the curves involved, they may be able to use the properties, etc., in the absence of perceptual support, and indeed, even in the presence of perceptual conflict. (2) The relations "as many as," "more than," and "fewer than," and their properties are basic in the development of the cardinal numbers. For this reason, an analogous study as suggested in (1) above is important. (3) If children are able to learn particular equivalence or order relations and their properties, relationships, and consequences, are they able to transfer this knowledge to other such relations given knowledge of that relation? (4) On a logical basis, the relations involved in this study are basic to measurement. Moreover, the relation of "more than," "fewer than," and "as many as " are basic to cardinal numbers. Are the relations basic also on a psychological basis?

Appendix

Sample Items of Measuring Instruments

Length Comparison Test

Material Set I

Materials:

One green stick; 3 pieces of white string, one being longer than, one shorter than, and one the same length as the green stick

Directions:

Item 1. Using these pieces of string, find a piece longer than this green stick.

Item 7. Using these pieces of string, find a piece shorter than this green stick.

Item 14. Using these pieces of string, find a piece the same length as this stick.

Conservation of Length Relations Test

Level I—Longer Than

Materials:

One green straw; 3 red straws, one being longer than, one shorter than, and one the same length as the green straw

Statement:

Using these red straws, find a straw longer than this green straw.

Transformation:

____________________ green

_______________________ red (move the red straw)

Question:

"Is this red straw still longer than this green straw?"

Level II--Longer Than

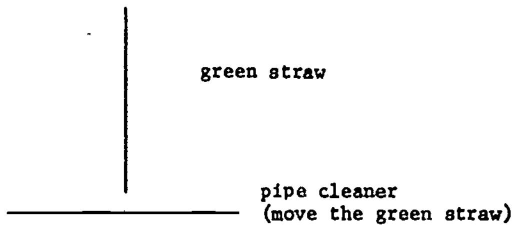
Materials:

One green straw; 3 white pipe cleaners, one being longer than, one shorter than, and one the same length as the green straw

Statement:

Using these pipe cleaners, find a pipe cleaner longer than this green straw.

Transformation:



Question: "Now is the green straw longer than the pipe cleaner?"

Reflexive and Nonreflexive Test

Materials:

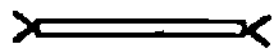
1 cardboard with M-L Diagram. 1 6-in. flannel strip

Statement:

Look at the length of this strip.

Transformation:

Look at the strip here. 

Look at the strip here. 

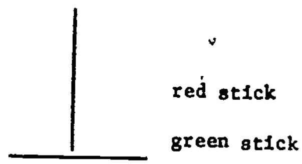
Question:

"Now, is the strip longer?"

Transitivity Test

Materials:

A red stick and a green stick of the same length attached to a cardboard as follows:



A white stick the same length as the red and green sticks for the child's use.

Question:

(a) "Is the red stick the same length as your stick?"
(b) "Is the green stick the same length as your stick?"
(c) "Is the green stick shorter than the red stick?"

DOUGLAS T. OWENS

Learning of Equivalence and Order Relations by Disadvantaged Five- and Six-Year-Old Children

Hilgard (1964, pp. 405-410) has suggested that there is a continuum of research studies along the dimension from pure research on learning to applied research on classroom practice. He specified six steps along this continuum: (1) learning research without regard for educational relevance, e.g., animal studies; (2) learning research using human subjects, but without concern for educational practice; (3) research on learning which is relevant to school learning, because school children and content are studied; (4) studies conducted in special laboratory classrooms on the feasibility of some educational practice; (5) tryout in a normal classroom; (6) developmental steps. In the first three of these steps the investigator is not primarily concerned with immediate application of his results to the classroom. In the second triad of steps, the researcher is expressly interested in classroom practice.

An analogous argument can be made for the existence of a continuum of types of research ranging from basic research on cognitive development to eventual classroom practices based upon cognitive development theory. From the Genevan studies it appears that Piaget and his colleagues are interested in the nature of cognitive development without particular concern for educational practice. Similarly, in many of the training studies which have been reported (Beilin, 1971), the experimenter is not primarily interested in developing curriculum for schools. The present study may be categorized at a level analogous to step (4) above. The investigator's goal is to make application of cognitive development theory to curriculum used in the classroom.

The experiment reported in this chapter is based on a doctoral dissertation in the Department of Mathematics Education at the University of Georgia (Owens, 1972).

The Problem

The purposes of this study were fourfold: (1) to determine the effectiveness of a set of activities designed to teach conservation and the transitive property of the matching relations "as many as," "more than," and "fewer than" to a group of economically disadvantaged five- and six-year-old children; (2) to determine the effect of the learning activities on the ability of the children to use properties of matching relations other than the specific properties upon which instruction was given; (3) to determine the effect of the learning activities on the ability of the children to conserve and use relational properties of length relations "as long as," "longer than," and "shorter than"; (4) to determine relationships among matching and length relations.

Operation and Structures

An operation, a concept central to Piaget's developmental theory (1970, pp. 21-23), has four properties. First, an operation is an action which can be carried out in thought as well as executed physically. The second characteristic of an operation is that it is reversible; the action can be carried out in one direction and in the opposite direction. Third, an operation always assumes some invariant (conservation). The fourth property is that every operation is related to a system of operation called a structure.

Piaget (Beth & Piaget, 1966, p. 172) believes that mental structures of the stage of concrete operations (from age 7 or 8 to age 12, approximately) may be reduced to a single model called "groupings." Piaget has postulated eight major groupings and a ninth preliminary grouping of equalities (Flavell, 1963, p. 195). If x and y represent grouping elements and "+" and "-" represent grouping operations, then each grouping has the following five properties (Piaget, 1964c, p. 42):

1. Combinativity, $x + x' = y$;
2. Reversibility, $y - x = x'$;
3. Associativity, $(x + x') + y' = x + (x' + y')$;
4. General operation of identity, $x - x = 0$;
5. Special identities, $x + x = x$.

In Grouping I, Primary Addition of Classes, the elements are classes which are ordered in a chain of inclusions $A \subset B \subset C$, etc. Addition, "+," is interpreted as the union of classes and "-" as set difference relative to a supraordinate class. Thus $A + A' = B$ where $A' = B - A$, since $A \subset B$, and "0" represents the null class (Flavell, 1963, pp. 173-74).

In Grouping V, Addition of Asymmetrical Relations, consider the seriation $0 < A < B < C < D$, etc. If $0 < A$, $0 < B$, $0 < C$, etc., are denoted by a, b, c, etc., and $A < B$, $B < C$, $C < D$, etc., are denoted by a', b', c', etc., respectively, then combinativity $(a + a' = b)$ is interpreted as transitivity of the relation when written as given (Beth & Piaget, 1966, p. 177).

In an additive system of relations, such as Grouping V, reversibility takes the form of what Piaget calls *reciprocity* (Beth & Piaget, 1966). Reciprocity consists of either permuting the terms of the relation (denoted by R), reversing the relation (R'), or both (R"). Thus, R(A < B) = B < A, R'(A < B) = A > B, and R" (A < B) = (B > A). Concerning reciprocity Beth and Piaget (1966) stated:

> *If we combine additively the relation (A < B) with its R, R', and R", we have:*
>
> *(1) (A < B) + (B < A) = (A = B) which is true in the case where the relation is ≤.*
>
> *(2) (A < B) + (A > B) = (A = B) id.*
>
> *(3) (A < B) + (B > A) = (A < B).*
>
> *Thus in all three cases there is no anullment, but the product is either an equivalence or the relation with which we started unchanged.* [p. 177]

Piaget indicated that statements (1) and (2) hold for partial order relations, such as "less than or equal to." Apparently (1) should be interpreted as, if A ≤ B and B ≤ A then A = B. This is precisely the antisymmetric property of a partial order relation. Moreover, if "<" is an order relation such as "less than" for example, then A < B and B < A contradict the asymmetric property and cannot hold simultaneously. In (3) A < B and B > A are logical equivalents and can hold simultaneously for order relations as well as for partial orderings. Thus, R" (A < B) = B > A is evidently the form of reciprocity characteristic of Grouping V since the R and R' cannot be combined with the original asymmetrical relation.

The general identity is not the absence of a relation, as in the case of the null class, but an equivalence relation. A special identity takes the form of (A < B) + (A < B) = (A < B), and associativity is limited to the cases in which no special identity is involved (Beth & Piaget, 1966, p. 178).

Development of the Concept of Measurement

Number

From Piaget's (1970, pp. 37-38) analysis of children's mental processes, he has concluded that the development of the concept of number is a synthesis of operations of class inclusion and operations of order. So long as the elements of a class have their qualities, Grouping I and Grouping V cannot be applied to the same elements simultaneously, but the basis of the notion of number is that the elements are stripped of their qualities, such that each element becomes a unit. As soon as the qualities of the elements are abstracted, Grouping I and Grouping V can no longer function separately but must necessarily merge into a single new structure (Beth & Piaget, 1966, pp. 259-67). "Class inclusion is involved in the sense that two is included in three, three is included in four, etc." (Piaget, 1970, p. 38). Since the elements are considered to be equivalent the only way to tell the elements apart is to introduce some order. The elements are arranged one after another spatially, temporally or in the counting sequence (Piaget, 1970, p. 38).

Van Engen (1971) disagrees with Piaget's notion of number.

> *The difficulty with this conception of number is that it does not distinguish between the elements of a set and the relation that exists between two or more elements of the set. The study of the order of whole numbers is the study of a relation that exists between two numbers and has the usual properties of an order relation.*[p. 40]

Van Engen (1971, pp. 37-39) suggests that, from a mathematical point of view, the cardinal numbers are standard sets of a particular kind. For example, 5 = {0, 1, 2, 3, 4}. To determine the cardinality of any set S, it is necessary to find one of the standard sets to which S is equivalent. This is accomplished by constructing a one-to-one correspondence or by counting. "From the point of view of mathematics, the relations 'as many as,' 'more than,' and 'fewer than' are basic to the development of number"(Van Engen, 1968, p. iii). On the basis of these relations, the cardinal numbers can be ordered, and the counting set can be formed.

These matching relations may be operationally defined between two sets A and B of physical objects as follows. Place an *a* beside a *b* until all the *a*'s or *b*'s are exhausted. If both sets are exhausted simultaneously, then there are as many *a*'s as *b*'s. If set B is exhausted and set A is not exhausted, there are more *a*'s than *b*'s and fewer *b*'s than *a*'s.

The relation "as many as" is thus another way of expressing set equivalence and is an equivalence relation. If "there are as many *a*'s as *b*'s" is indicated by "$A \approx B$" for equivalent sets A and B, then "$\approx$" is reflexive ($A \approx A$); symmetric (If $A \approx B$ then $B \approx A$); and transitive (If $A \approx B$ and $B \approx C$ then $A \approx C$). The relations "more than" and "fewer than" are order relations, for if $A > B$ indicates "there are more (or fewer) a's than b's," then ">" is nonreflexive ($A \not> A$); asymmetric (If $A > B$ then $B \not> A$); and transitive (If $A > B$ and $B > C$ then $A > C$). The relations "more than" and "fewer than" are examples of asymmetrical transitive relations of which Piaget wrote. They also exhibit the reversibility property. For if there are more *a*'s than *b*'s, then there are fewer *b*'s than *a*'s , and conversely. Thus, from the mathematical point of view of Van Engen and from the psychological perspective of Piaget, the matching relations are involved in the development of the number concept.

Measurement

Measurement has been described as "a process whereby a number is assigned to some object"(Steffe, 1971, p. 335). From this definition it follows logically that number is a prerequisite of length. Sinclair (1971) has stated that the

> *first measurement concept (length) is achieved rather later than that of number;. . . There is an even greater time lag... between acquisition of the corresponding conservation of length concept and the simple numerical conservations. Although the psychological construction is parallel, dealing*

> *with continuous elements is very much more difficult than dealing with discontinuous units.* [p. 153]

Sinclair (1971) has presented empirical evidence which is consistent with the logical conclusion that number precedes length in development. However, the explanation given is that length is achieved later than number because it is more difficult to deal with continuous elements than with discrete objects. Relations also provide a basis for the development of measurement in elementary school children. The relations "as long as," "longer than," and "shorter than" are comparisons of the relative lengths of segments, but for children they can be defined on objects such as sticks or straws. In the present study, a concept of number was not required in establishment of either the matching relations or length relations. However, the materials for the length relations were continuous objects, and the materials for the matching relations were discrete. Thus, Sinclair's (1971) hypothesis raises the question of whether the ability to use the matching relations precedes the ability to use length relations in the tasks of conservation and transitivity.

For an operational definition of the length relations, consider two segments A and B. A is as long as B if whenever A and B

> *...(or their transforms) lie on a line in such a way that two end points coincide (right or left), the two remaining end points coincide. A is longer than B if and only if the remaining end point of B coincides with a point between the end points of A. Also, in this case, B is shorter than A.* [*Carey and Steffe, 1968*, p. 31]

The relation "as long as" thus defined is an equivalence relation and has the reflexive, symmetric, and transitive properties as does the matching relation, "as many as." The relations "longer than" and "shorter than" are order relations and possess the nonreflexive, asymmetric, and transitive properties analogous to the relation "more than."

Conservation and Transitivity

In Piaget's (1952) classical conservation of number tasks, a child is asked to establish that there are as many objects in a set A as in set B. Then one of the collections, say A, is taken through a physical transformation. Then the child is asked, "Are there as many *a*'s as *b*'s, or does one have more?" Van Engen (1971, p. 43) has argued that this task may be measuring whether or not the child conserves the one-to-one correspondence rather than conservation of number. In this study a task similar to the above example is considered to be a measure of conservation of the relation "as many as." It is not necessary that conservation be limited to cases of equivalence. For example, in a task given by Smedslund (1963b) a child was asked to establish that one stick was longer than a second stick and to maintain that the one stick was longer after a conflicting cue was introduced. While Smedslund called the task "conservation of length," a similar task in the present study is called "conservation of the relation 'longer than.'" Thus, order relation conservation is also included.

Conservation is studied from the relational point of view and transitivity is necessarily a relational property. Thus, the relationship between the development of conservation and attainment of transitivity is approached from the standpoint of relations. In his earlier writing, Piaget (1952, p. 205) reported that as soon as children can establish a lasting equivalence (that is, conserve the equivalence), they can at once use the transitive property. "The explanation is simple: the composition of two equivalences [transitivity]* is already implied in the construction of a single lasting equivalence between two sets, since the different successive forms of the two sets seem to the child to be different sets" (Piaget, 1952, p. 208).

Similarly, Northman and Gruen (1970) argue that transitivity is involved in equivalence conservation. Suppose the subject establishes A equivalent to B ($A = B$). When an equivalence-preserving transformation T is performed, the subject establishes (covertly) $A = T(A)$. Then, transitivity is used in order to deduct $T(A) = B$ or to conserve the equivalence of A and B.

Smedslund (1964) has argued that from a logical point of view, conservation precedes transitivity in the child's development. Consider three quantities which are related by a transitive relation @. Assume that a child established A@B. B (or A) must undergo some transformation, T, before B is compared with C; otherwise, A and C can be compared perceptually. Hence $B = T(B)$ (or $A = T(A)$) must hold from one comparison to the other.

In a later discussion of training research Piaget (Beth & Piaget, 1966, p. 192) also alluded to an ordering in the attainment of conservation and transitivity. He reported that Smedslund easily induced conservation of weight by repeatedly changing the shape of a small clay ball and checking the weight on a scale. Smedslund was not successful in obtaining immediate learning of the transitive property.

Basic Questions of the Study

Among equivalence and order relations in the primary school curricula are set relations, based on matching finite sets of objects, and length relations, determined by comparing relative lengths of objects. It appears from Piaget's theory of grouping structures that if a child has, for example, Grouping V: Addition of Asymmetrical Relations in his cognitive structure, he can use logical properties of any such relations. In this study, an attempt was made to: (1) provide experiences for five- and six-year-old children in which the transitive property of the relations "as many as," "more than," and "fewer than" could be observed empirically by the children, and in which conservation of these relations could be observed through reversing a transformation; and (2) determine the effectiveness of the treatment in inducing the logical-mathematical properties of these relations and of the length relations "as long as," "longer than," and "shorter than."

Specifically, the following questions are basic to the study where most, but not all, arise from Piaget's theory.

---

*Added by the Author.

1. What is the effect of selected experiences on the ability of children to establish, conserve, and use properties of equivalence and order relations?

2. What is the effect of age on the ability of children to establish, conserve, and use properties of equivalence and order relations?

3. To what extent does an experimentally induced capability to conserve and use transitivity of matching relations transfer across relational categories to conservation and transitivity of length relations?

4. To what extent does an experimentally induced capability to conserve and use transitivity of matching relations transfer to remaining properties of matching relations?

5. What is the effect of a pretest on the ability of children to establish, conserve, and use properties of relations with or without selected experiences?

6. Is the ability to conserve matching relations related to the ability to use the transitive property of matching relations?

7. Is the ability to conserve length relations related to the ability to use the transitive property of the length relations?

8. Is the ability to conserve matching relations related to the ability to conserve length relations?

9. Is the ability to use transitivity of matching relations related to the ability to use transitivity of length relations?

10. Is the ability to solve a problem involving transitivity of a matching relation related to performance on a test of conservation or transitivity of matching relations which utilizes a standardized interview technique?

11. What are the intercorrelations among the variables of the study?

Method

The subjects of the study were 23 kindergarten and 24 first-grade children of the William Fountain Elementary School, Atlanta, Georgia. Kindergarten children were randomly selected from 35 children of two classes whose ages were in the range (5:1)[1] to (5:10) at the outset of the study. Grade one children were randomly chosen from 48 children of three classes with ages between (6:1) and (6:10) at the outset. The school was chiefly composed of Negro children from low income families. With one exception, the children in the sample were Negro.

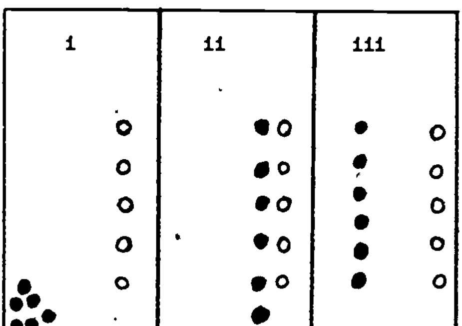
[1]5 years, 1 month.

Tests

The thirteen tests described below were constructed to measure the abilities of the children to establish relations, conserve relations, and use relational properties.

The *Matching Relations (MR) Test* was designed to measure the ability of a child to establish matching relations and the *Conservation of Matching Relations (CMR) Test* was designed to measure the ability of a child to conserve a matching relation, provided that he could establish the relation. These two tests were administered simultaneously. In the example presented in Figure 1, a child was given five blue discs glued on a piece of cardboard and six red discs (i). He was instructed to pair the red discs and the blue discs. After the pairing (Figure 1—ii), the examiner asked two questions, "Are there as many red discs as blue discs?" and "Are there more red discs than blue discs?" After the second response the examiner rearranged the red discs (Figure 1-iii) and repeated the same two questions. In each case the correct answer to one question was "yes" and to the other "no." In each item, the rearrangement was perceptually biased in favor of the incorrect conclusion. The first two questions comprised an item of the MR Test. All four questions were considered in the CMR Test.

Figure 1



The *Length Relations (LR) Test* was designed to measure the child's ability to establish length relations. The *Conservation of Length Relations (CLR) Test* was designed to measure the ability of a child to conserve length relations. These two tests were given together in the same way as the MR and CMR Tests. In each item the child was asked to establish a length relation between two sticks (or straws) by answering two questions. Then the sticks were rearranged to produce a perceptual bias against the correct conclusion, and the questions were repeated.

The purpose of the *Transitivity of Matching Relations (TMR) Test* was to measure a child's ability to use the transitive property of matching relations. On a TMR item a child was presented three collections A, B, C, of physical materials, arranged in clusters. Suppose, for example, that there were fewer *a*'s than *b*'s and fewer *b*'s than *c*'s. The child was instructed to pair the *a*'s and *b*'s and was then asked, "Are there fewer *a*'s than *b*'s?" The examiner then put the *a*'s into a cup which sat nearby and said "Pair the *b*'s and *c*'s." After the pairing the examiner asked, "Are there fewer *b*'s than *c*'s?" The examiner then placed the *c*'s in another cup and asked, "Are there fewer *a*'s than *c*'s?" and "Are there more *a*'s than *c*'s?" (or "Are there as many *a*'s as *c*'s?") Note that the sets A and C were not "paired" and that the objects were screened at the time of the transitive inference.

The *Transitivity of Length Relations (TLR) Test* was designed to measure the ability of a child to use the transitive property of the length relations. On each item, as in the TMR test, a child was asked to establish the relation between two sticks, A and B. Stick A was placed in a box and stick B was compared with another stick C such that the same relation held between B and C as between A and B. Then stick C was placed in a box and two questions, relative to A and C, were asked.

The purpose of the *Symmetric Property of the Matching Relations (SMR) Test* was to determine the child's ability to use symmetry of the relation "as many as." For an item of SMR test the child was presented two collections A and B of objects and instructed to pair the objects. After the pairing the examiner asked two questions: "Are there as many *a*'s as *b*'s?" (Response), "Are there more (or fewer) *a*'s than *b*'s?" (Response). Then the examiner put the two collections into two cups and asked, "Are there as many *b*'s as *a*'s?" (Response), and "Are there more (or fewer) *a*'s than *b*'s?".

The *Symmetric Property of the Length Relations (SLR) Test* was designed to measure the ability of a child to use symmetry of the relation "as long as." In this case the child was asked to compare two sticks (straws) to determine that stick A was as long as stick B and that an order relation did not hold. Then the examiner placed the two sticks into two boxes and asked, "Is stick B as long as stick A?" and the previous question involving an order relation.

The *Test of the Asymmetric Property of the Matching Relations (AMR)* was designed to measure the ability of a child to use the asymmetric property of the relations "more than" and "fewer than." The child was presented two collections, for example, with more *a*'s than *b*'s and instructed to pair them. After the pairing the examiner asked, "Are there more *a*'s than *b*'s?" After the response the examiner placed the two collections into two cups and asked, "Are there more *b*'s than *a*'s?" and "Are there more *a*'s than *b*'s?"

The purpose of the *Test of the Asymmetric Property of the Length Relations (ALR)* was to measure the child's ability to use the asymmetric property of "longer than" and "shorter than." On an item of this test the child compared two sticks related by an order relation. Suppose a child established correctly that stick A was shorter than stick B. The examiner then placed each stick into a box and asked, for example, "Is stick B shorter than stick A?" and "Is stick A shorter than stick B?"

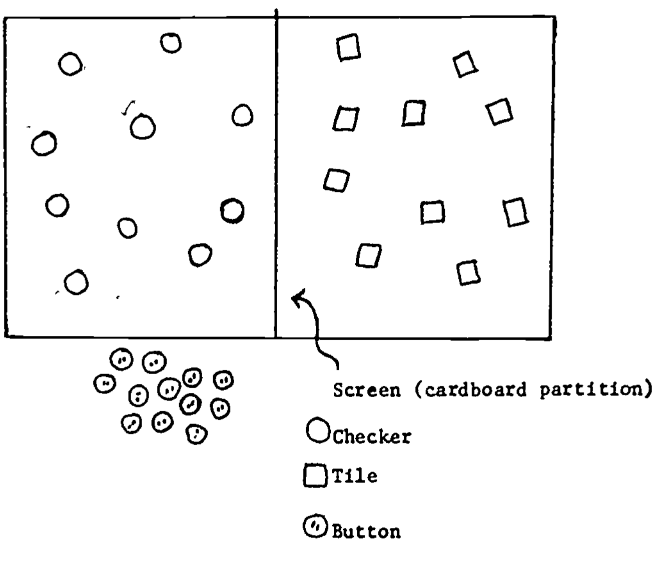
The *Reversibility of Matching Relations (RMR) Test* was designed to measure the child's ability to use the following property: if there are more (fewer) *a*'s than *b*'s, then there are fewer (more) *b*'s than *a*'s. On a given item, the child was presented with two collections A and B of objects such that an order relation held. After the child had paired the objects, the examiner asked, "Are there more (fewer) *a*'s than *b*'s?" The examiner then put the objects into two cups and asked, "Are there fewer (more) *b*'s than *a*'s?" and "Are there as many *b*'s as *a*'s?"

The *Reversibility of Length Relations (RLR) Test* was designed to measure the ability of a child to use the following reversibility property: if segment A is longer (shorter) than segment B, then segment B is shorter (longer) than segment A. After the child compared two sticks, the examiner asked, "Is stick A longer (shorter) than stick B?" Then, the examiner placed the sticks into two boxes and asked, "Is stick B shorter (longer) than stick A?" and "Is stick B as long as stick A?"

On each of the MR, LR, CMR, CLR, TMR, and TLR Tests there were two items which in fact exhibited each relation. Thus there were six items on each of these tests. The SMR and SLR Tests contained three items each. The AMR, ALR, RMR, and RLR Tests had two items for each of two order relations, or four items per test. The total number of items was 58. The relations (MR, LR) Tests involved situations under which the stimuli were arranged to aid the child in establishing the relation. The questions of the conservation (CMR, CLR) Tests were administered under conditions of perceptual conflict. All other items were administered under screened stimuli conditions. The child was not asked to give reasons for his answers on any of the structured items.

The *Transitivity Problem (TP)* was designed to measure the ability of a child to solve a problem which involved transitivity of a matching relation with minimum guidance from the examiner. The situation involved a cardboard box from which the front and top were removed. The box was divided into halves by a partition as shown in Figure 2. Ten checkers were attached to the bottom inside one half of the box and ten tiles were attached in the other side. Twelve buttons lay on the table in front of the box. After the objects were identified, the examiner said, "Find out if there are as many checkers as tiles. You may use the buttons to help you find out." In general the examiner gave as little guidance as was possible, but if the child failed to respond at some point, the examiner directed the next step toward solution. When a response was given, the examiner asked for an explanation.

Figure 2



Scoring Tests

An item was scored "pass" provided that a child answered correctly all the questions contained in the item and "fail" otherwise. The number of items scored "pass" by a child on each test was considered to be his score on the test. For the purpose of comparing these data with other studies it was desirable to distinguish children for which evidence existed that they could use a property from those for which no such evidence existed. This was accomplished by setting a criterion score based on a random model. It was assumed that a child could use a relational property if and only if he met the criterion on a particular test. Four of the six items was the criterion set on each of the CMR, CLR, TMR, and TLR Tests. The probability of reaching this criterion by guessing was at most .038.

For the Transitivity Problem, the following four levels of ability to apply the transitive property were identified: 1—the child neither consistently established relations nor used the transitive property; 2—the child established relations but did not use the transitive property; 3—the child both established relations and used the transitive property without adequate justification; 4—the child established relations, used transitivity, and gave adequate justification for his conclusion. The consensus of two of three judges' ratings, based on transcripts of audio tapes, was taken as the child's rating on the Transitivity Problem.

Instructional Activities

All of the instructional activities were designed for use in small instructional groups and involved manipulative materials. In some activities each child had his own set of materials. Other activities involved one set of materials for the entire group. In the latter cases, the instructor or one child performed the manipulations, but all of the children entered into discussion. Materials for instruction varied from materials such as small toys to neutral material such as checkers, tiles, or colored wooden discs. Colored sticks, straws, etc., represented segments for length comparison.

The purpose of Unit I, *Matching Relations*, was to develop the ability of the children to establish matching relations. The relations were introduced by having children pair the objects from two finite sets. It was noted that the sets may or may not be in one-to-one correspondence. When the sets A and B were equivalent, the phraseology "there are as many a's as b's" was used. "More than" was introduced second, and "fewer than" was introduced as the reverse of "more than." It was emphasized that if a relation holds between two sets (in a fixed order) then no other relation holds.

Unit II, *Length Relations*, was designed to develop the ability of children to establish length relations. The relations were introduced by placing the ends of two sticks together, observing the remaining ends, and associating the name of the appropriate relation. After "longer than" was discussed, "shorter than" was introduced as the reverse. The equivalence relation "as long as" was the third length relation considered.

The purpose of Unit III, *Conservation of Matching Relations*, was to develop the ability of children to maintain relations between sets when the physical matching of the objects is destroyed. The principle of reversibility of a transformation was emphasized by having the children return the objects, following a transformation, to the position in which the relation was established. Combinations of perceptual screening, perceptual conflict, child transformations of his own materials, and instructor transformations in a group situation were used in Unit III and Unit IV, *Transitivity of Matching Relations.* Unit IV was designed to develop the ability of children to use the transitive property of the matching relations. The chief method of the transitivity training was what has been termed fixed practice with empirical control (Smedslund, 1963a). The instructor gave explicit instructions for comparing sets A and B, then B and C. Sets A and C were compared after the child made a prediction of the relation between them.

Design

Age (five-and six-year-olds) was used as a categorization variable because of its importance in cognitive development. Treatment was a second major factor of which two levels existed, a Full Treatment and a Partial Treatment. The Full Treatment consisted of all four units described earlier and the Partial Treatment consisted of Units I and II. Transfer shall be inferred from a significant difference in favor of the full treatment group in performance on some property for which no instruction was given, provided that there is a significance in the same direction on a related property for which instruction was given.

In most of the learning research based on cognitive development theory, and in many other educational research studies, pretests are given to all subjects. However, due to the large time requirement of instructing and testing, a strictly pretest-posttest design was not considered feasible in the present study. However, it was desirable to obtain premeasures on some subjects, so a Solomon four group design was selected. Use of this design requires that part of the Full Treatment group and part of the Partial Treatment group be selected at random to have the pretest. Schematically, the Solomon four group design may be represented as follows:

Randomized assignment	G_1	Pretest:	Full Treatment	Posttest
to groups	G_2	Pretest:	Partial Treatment	Posttest
	G_3		Full Treatment	Posttest
	G_4		Partial Treatment	Posttest

Campbell and Stanley (1963) note that the results of an experiment using this design are more generalizable than those from a pretest-posttest design, because the effects of testing and the interaction of testing with the treatment are determinable, and randomization controls for initial biases between groups. They also suggest that experience with the Solomon four group design in a particular research area gives information about the general likelihood of the effects of pretesting in that area of research. It is desirable that this general information be obtained for learning research based on cognitive development theory.

Campbell and Stanley (1963) suggest that to analyze the data obtained from the Solomon four group design, the pretests may be disregarded except as another treatment, and only the posttest scores analyzed by analysis of variance. If there is no effect due to pretesting or no interaction of pretesting and treatment, then the pretest data may be used as a covariate in performing analysis of covariance on the data from the pretest groups. In the present study the first of these suggestions was followed, but the analysis of covariance was not performed due to the small pretest group size. Moreover, the pretest data were considered to be somewhat invalid because the examiners strictly adhered to the relational terminology preferred by the investigator. A more flexible testing procedure, adaptable to the child's language pattern, was used for the posttests.

Procedure

Children in the Full and Partial Treatment groups first had experience in establishing relations (Units I and II). Ten lessons on matching relations from Unit I and seven lessons from Unit II on length relations were given. Then the tests on relations (MR, LR), conservation (CMR, CLR), and transitivity (TMR, TLR) of each relational category were administered as pretests to the pretest group while the no-pretest group had only the relations tests. Following the pretests, the Full Treatment group had four lessons on conservation of matching relations (Unit III) and five lessons on the transitive property (Unit IV). Near the end of this instructional period the Partial Treatment group had two additional lessons on matching relations, but the remainder of the treatment period was spent in normal classroom activities.

Each lesson was of 20-30 minutes duration. There were four to six children in an instructional group. The investigator and two teachers' aides served as instructors and testers. Instructional groups were rotated among instructors each day. During testing the Full and Partial Treatment groups were balanced among testers in five untimed interviews per child, except that the Transitivity Problem session (sixth interview) was held entirely by the investigator. The test items given during a test session were randomly ordered for each child, independently of other children, and each pair of test questions of an item were randomly ordered for each item and each child.

Near the end of the study it was apparent that the full treatment had not extensively changed the language patterns of the children with regard to relational terminology. The investigator felt that strict adherence to predetermined terminology could make the tests invalid in terms of the concepts measured. Approximations to desired terminology, for example, "the same" for "as many as," was accepted in the posttests. Further, if a child were giving a "no--no" or a "yes--yes" response set to an item, the question was repeated using an alternate terminology. This was the only way in which the postests differed from those tests which were given as pretests, but this was considered to be sufficient to make the pretest data invalid per se. Thus, Pretest was retained as a factor but the data were disregarded.

Statistical Analyses

The data for the analysis of variance were vectors of 12 posttest scores for each individual. Multivariate analysis of variance was an appropriate statistical design. The design is represented diagramatically in Table 1. Three factors were considered at two levels each: Age--

five- and six-year-olds; Treatment--full treatment and partial treatment; and Pretest--pretest and no-pretest. The Age X Treatment interaction and the Treatment X Pretest interaction were of particular interest to the investigator, but the three-way Age X Treatment X Pretest interaction was not considered as one of the questions in the study. Thus, three separate multivariate analyses of variance were performed in which two factors and their two-way interaction were considered. Data for all cells of each factor were combined for the 2 X 2 factorial multivariate analyses. It was of particular interest to the investigator to determine the effects of the treatment, the pretest, and age upon each of the variables and to determine if two-way interactions existed. Thus, a univariate analysis for each of the 12 response variables was performed. In this regard Table 1 may be interpreted as 12 univariate designs in which the three main effects and pairwise interactions were of interest. Again the analyses were handled by consideration of three 2 X 2 factorial univariate analyses for each of the 12 variables and combining all cells within a factor under consideration.

Table 1

Diagram of the Design

Factors and Levels		Twelve Response Variables											
Treatment	Age	MR	CMR	TMR	SMR	AMR	RMR	LR	CLR	TLR	SLR	ALR	RLR
Full Treatment	Five												
	Six												
Partial Treatment	Five												
	Six												
Treatment	Pretest												
Full Treatment	Pretest												
	No Pretest												
Partial Treatment	Pretest												
	No Pretest												

Calculations for all of the MANOVA's and ANOVA's were performed by computer with the use of the computer program MUDAID (Applebaum and Bargmann, 1967). MUDAID provides multivariate and univariate analyses of variance for pairs of factors and pairwise interactions. Also, each multivariate pass provides matrices of intercorrelations among the response variables.

Each covariance matrix in a multivariate analysis contains estimates of the variances of the variables on the main diagonal and estimates of the covariances for pairs of variables in the off-diagonal positions. Each covariance matrix has an associated matrix of sums of squares and cross products. The sum of squares of error and sums of products of error are the residuals after the effects of the factors and interactions have been removed by subtraction of their sums of squares and sums of products from the respective totals. The correlations reported in this study were calculated from the covariance matrix derived from the matrix of sums of squares and products of error in the Treatment X Age analysis.

Chi-square tests for independence (Ferguson, 1966, pp. 192-208) were used to determine whether a relationship existed between levels of performance on the Transitivity Problem and Age. Chi-square tests were also made to determine relationships between conservation and transitivity within a relational category and to determine relationships across relational categories within corresponding measures. Chi-squares were calculated on the 2 X 2 or 2 X 3 tables where the frequencies were the number of children achieving a criterion or level of performance.

Results

Multivariate Analyses

None of the *F* ratios for any factor or two-way interaction were significant at the .05 level of significance in the multivariate tests. However, the *F* statistic for the main effect of Age was 1.95 in the Treatment versus Age multivariate analysis with 12 and 32 *df*. The critical value ($p < .05$) of *F* with 12 and 32 *df* is 2.07. Thus the factor Age approached significance, but no interpretation was made.

Univariate Analyses

Analyses of Variance for which *F* ratios were significant in the Treatment versus Age analyses are reported in Table 2. Table 3 contains analyses of variance for the cases of significance in the Treatment versus Pretest analyses. Any factor which was statistically significant in a Pretest X Age analysis was significant in the corresponding Treatment X Pretest or Age X Pretest analysis. Thus, analysis of variance tables are not presented for Pretest versus Age. Group means, as percents, for treatment and age groups are presented in Table 4. Age was the only significant ($p < .01$) effect for the variables matching relations (MR) and conservation of matching relations (CMR). In the first case, the mean for the six-year-olds was 87% and for the five-year-olds was 59%. On conservation the six-year-old group performed at a mean of 62% and the five-year-old group at a mean of 36%. It was not anticipated that Treatment would be significant for MR since all children had received instruction in matching relations.

Table 2
Treatment Versus Age Analyses of
Variance With Significant Effects

Source of Variation	Relations (MR) M.S.	Relations (MR) F	Conservation Of MR (CMR) M.S.	Conservation Of MR (CMR) F	Transitivity Of MR (TMR) M.S.	Transitivity Of MR (TMR) F	Symmetric Property Of MR (SMR) M.S.	Symmetric Property Of MR (SMR) F
Treatment (T)	.44	.16	2.19	.62	19.90	8.34**	2.32	2.69
Age (A)	31.53	11.79**	28.88	8.18**	2.36	.99	11.10	12.89**
T X A	.15	.06	1.49	.42	.16	.07	.15	.18
Error	2.67		3.53		2.38		.86	

	Asymmetric Property of MR (AMR) M.S.	Asymmetric Property of MR (AMR) F	Reversibility of MR (RMR) M.S.	Reversibility of MR (RMR) F	Length Relations (LR) M.S.	Length Relations (LR) F
Treatment	7.69	3.98√	2.60	1.97	1.00	1.01
Age	10.39	5.37	8.74	4.79*	5.87	5.89*
T X A	.14	.07	.47	.26	.53	.53
Error	1.93		1.83		1.00	

√(p < .10), *(p < .05), **(p < .01)
Note: Each factor and interaction had 1 *df*; error 43 *df*.

Table 3
Treatment versus Pretest Analyses of
Variance with Significant Effects

	Transitivity of MR (TMR) M.S.	Transitivity of MR (TMR) F	Asymmetric Property of MR (AMR) M.S.	Asymmetric Property of MR (AMR) F	Symmetric Property of LR (SLR) M.S.	Symmetric Property of LR (SLR) F	Reversibility of LR (RLR) M.S.	Reversibility of LR (RLR) F
Treatment (T)	20.41	8.45**	8.37	4.10*	.21	.18	1.71	1.21
Pretest (P)	1.19	.49	5.39	2.64	4.71	4.09*	3.53	2.31
T X P	.00	.00	.39	.19	4.92	4.29	9.54	6.26*
Error	2.42		2.04		1.15		1.52	

*p < .05 **p < .01
Note: Each factor and interaction had 1 *df*; error 43 *df*.

Table 4
Group and Total Means, as Percents, for Each of 12 Variables

Variable	Treatment Groups		Age Groups		Totals
	Full	Partial	Six	Five	
MR	75	71	87	59*	73
CMR	53	45	62	36*	49
TMR	58	36*	51	43	47
SMR	81	48	81	48*	65
AMR	73	52*	74	50*	62
RMR	68	54	72	50*	61
LR	90	85	93	81*	87
CLR	62	52	63	51	57
TLR	47	51	50	48	49
SLR	62	58	67	54	60
ALR	64	57	68	53	61
RLR	72	62	65	70	67

*This pair of means was significantly different in a univariate analysis.

Treatment was a significant ($p < .01$) main effect for transitivity of matching relations (TMR). The full treatment group mean was 58% and the partial treatment group mean was 36%. Treatment was also a significant ($p < .05$) factor for the variable AMR in the Treatment versus Pretest analysis, and was close to significance at the .05 level in the Treatment versus Age analysis. In this case the means were 73% and 52% for the full treatment and partial treatment groups, respectively. Age was also a significant ($p < .05$) main effect for AMR as it was for SMR ($p < .01$) and RMR ($p < .05$). In each of these cases the six-year-olds performed at a higher level than the five-year-olds.

The *F* statistic for the factor Age and for the variable Length Relations (LR) was significant. However, in Bartlett's test (Ostle, 1963, pp. 136-137) the hypothesis of homogeneity of variances was rejected. Thus, no interpretation of the ANOVA was made.

There were no significant interactions in the Treatment versus Age analyses. There were, however, two Pretest X Treatment interactions ($p < .05$) for the variables in SLR and RLR. Pretest was not a significant main effect in the absence of interaction in any analysis. The cell means for the significant interactions are presented in Table 5. In each case the greatest mean was that of the full treatment group which had no pretest, and the least mean was that of full treatment group which had pretests. One possible interpretation of this interaction is that the pretests interfered with the effect of the treatment. However, this may be a misinterpretation since instruction was not given on the symmetric and reversibility properties of either category of relations, nor was there any indication of transfer to the properties of length relations from the instruction which was given. The interpretation which is accepted here is that the pretests had essentially no effect on the subjects' performance on the posttests.

Table 5
Cell Means as Percents: Treatment X Test Interactions

	SLR			RLR		
	Pretest	No Pretest	Total	Pretest	No Pretest	Total
Full Treatment	42	85	62	54	91	72
Partial Treatment	58	58	58	67	58	63
Total	50	71	60	60	74	67

Age was the most general effect in the study, but the surprising result was that Age was not significant for any length relational variable. In comparing means for length relational variables with means for matching relational variables, it may be noted that a grand mean of 87% for the variable Length Relations was equal to the mean for the six-year-old group on MR, which was significantly greater than the mean for the five-year-old group on MR. The grand mean of 57% for CLR was between the means of 62% for the six-year-old group and 36% for the five-year-old group on CMR. For TLR the mean of 49% was between the significantly different means for the full and partial treatment groups for TMR. The means of 60%, 61%, and 67% for SLR, ALR, and RLR, respectively, were between the respective matching relational means for the two ages, which were different because of an age effect in each case. Also, in each case the mean for the five-year-old group was greater for length than for the corresponding matching relations variable. Thus, while no factors were significant for the length relations variables, overall performance in each case was not decidedly different from performance on the corresponding matching relations variable. No formal statistical tests were made between variables across relational categories.

Transitivity Problem Results

In order to test the relationship between performance on the Transitivity Problem and the factors Treatment and Age, chi-square tests for independence were performed on contingency tables. The frequencies of Transitivity Problem ratings versus Treatment groups are presented in Table 6, and ratings by age frequencies are found in Table 7. Two children counted and no ratings were possible. While it is of interest to see the number of children at each of the four levels on the Transitivity Problem, categories 3 and 4 were combined into a single category, 3 or 4 (the child used transitivity), for the chi-square tests. This was necessary to increase the expected frequency for some cells. Frequencies are presented both ways but the chi-square tests were performed on the 2 X 3 tables.

Table 6
Contingency Table: Transitivity Problem
Ratings Versus Treatment Group

Treatment Group	Rating				
	1	2	3 or 4	3	4
Full Treatment	4	7	10	6	4
Partial Treatment	7	12	5	4	1

Table 7
Contingency Table: Transitivity Problem
Ratings Versus Age Level

Age Level	Rating				
	1	2	3 or 4	3	4
Six	3	7	12	8	4
Five	8	12	3	2	1

The chi-square calculated for Table 6 was 3.62 with 2 *df*, not significant at the .05 level. Thus while there appeared to be a tendency for more children in the full treatment group to get a rating of 3 or 4 and more children in the partial treatment group to get a rating of 1 or a rating of 2, the hypothesis of independence was not rejected. The chi-square calculated for Table 7 (2 X 3) was 8.97 with 2 *df* which is significant at the .02 level. Thus, the null hypothesis of independence was rejected, and the existence of a relationship between age and the level of performance on the Transitivity Problem was accepted. There was a tendency for six-year-old children to have the higher rating of 3 or 4, and for the five-year-old children to have the lower ratings of 1 and 2.

While the treatment was effective in improving the abilities of the children to perform the transitivity tasks of TMR, the treatment was not related to level of performance on the Transitivity Problem. On the other hand, there was no significant difference between ages in performance on TMR, but age level was related to level of performance on the Transitivity Problem. These results raise a question about the relationship between performance on the Transitivity Problem and the more structured tests.

Relationships Among the Variables

Chi-square tests were used to test for a relationship between level of performance on the Transitivity Problem and criterion performance on TMR and CMR. The frequencies of the ratings on the Transitivity Problem versus meeting the criterion on TMR and CMR are presented in Table 8 and Table 9 respectively. Chi-square tests were run on the 2x3 tables.

Table 8
Contingency Table: Ratings on Transitivity Problem Versus Criterion on Transitivity of Matching Relations

TMR Criterion Level	Rating 1	2	3 or 4	3	4
Criterion	1	7	8	5	3
Not Criterion	10	12	7	5	2

Table 9
Contingency Table: Ratings on Transitivity Problem Versus Criterion on Conservation of Matching Relations

CMR Criterion Level	Rating 1	2	3 or 4	3	4
Criterion	0	5	13	8	5
Not Criterion	11	14	2	2	0

The chi-square calculated for Table 8 was 5.45. The critical value of chi-square with 2 *df* is 5.99 ($p < .05$). Thus, the chi-square for level of performance on the Transitivity Problem versus Transitivity, as measured by TMR test, was near significance at the .05 level, but independence was accepted. The chi-square calculated for Table 9 was 22.43 ($p < .001$). There was a strong relationship between ratings on the Transitivity Problem and achieving the criterion on the CMR test.

The product moment correlations in the present study were calculated in the multivariate analysis by using the error covariance matrix from the Treatment versus Age Analysis. The reason for using this error matrix to calculate the correlations is that essentially all significant effects have been eliminated from the matrix and only nonsignificant effects remain. That is, the effects of Treatment and Age were statistically removed by subtraction, and only the (nonsignificant) effects of Pretest remain. The correlations are presented in Table 10. Since *df* for error in the analyses of variance was 43, there are 42 *df* associated with each correlation of Table 10. The critical values for correlations significantly different from zero are .30 ($p < .05$) and .39 ($p < .01$).

Inspection of Table 10 revealed that 47 of the 66 correlations were significantly different from zero and all were positive. Only two correlations were greater than .60 and 16 others were greater than .50. Of the 19 nonsignificant correlations, 13 were with or between LR and TLR. It was interesting that the only length variable with which LR was correlated was CLR. Indeed, each item of the CLR Tests was dependent upon an item of the LR Test. It appears that there is little relationship between each of LR and TLR and the remaining variables. In addition to TLR, three variables are not correlated with CLR. The nonsignificant correlations of SMR with RMR and RLR indicate a lack of relationship between the symmetric property of "as many as" and the reversibility property of either relational category. The additional nonsignificant correlation

Table 10
Intercorrelations Among the 12 Variables

	CMR	TMR	SMR	AMR	RMR	LR	CLR	TLR	SLR	ALR	RLR
MR	73**	57**	32*	59**	59**	22	35*	16	49**	55**	40**
CMR		51**	37*	59**	48**	30*	41**	40**	56**	48**	39**
TMR			44**	28	47**	07	24	28	51**	59**	32*
SMR				46**	27	43**	41**	36*	55**	51**	17
AMR					45**	43**	40**	26	54**	52**	41**
RMR						07	27	25	54**	46**	50**
LR							46**	15	27	24	20
CLR								11	53**	25	39**
TLR									47**	33*	23
SLR										64**	57**
ALR											40**

*p < .05, **p < .01
Note: Decimal points are omitted.

was between TMR and AMR. The remaining correlations with each matching relational variable were significant. It is interesting to note that CMR was correlated with each variable across both relational categories.

Whether or not a child in the present study attained the criterion on a particular test is a measure of the child's ability to use the relational property of the test. In order to examine the hypothesis that conservation ability precedes the ability to use the transitive property within a category of relations, 2 x 2 frequency tables, of those who did and did not use conservation and transitivity, were prepared. Chi-square tests for independence were then made on the contingency tables. The frequencies of children meeting criterion on CMR versus meeting criterion on TMR are presented in Table 11. Table 12 contains the frequencies of children in the sample who met the criterion on conservation versus those who met the criterion on transitivity of length relations. The calculated chi-squares with 1 *df* were 1.73 for Table 11 and .57 for Table 12. These nonsignificant chi-squares indicate independence between the ability to use conservation and the ability to use transitivity within the respective relational categories. These results are not completely consistent with the significant product moment correlation of .51 between CMR and TMR for the matching relations. However, in the case of length relations, the result is consistent with the nonsignificant correlation between conservation and transitivity.

Table 11
Contingency Table: Criterion on CMR Versus Criterion on TMR

Conservation of Matching Relations (CMR)	Transitivity of Matching Relations (TMR)	
	Criterion	Not Criterion
Criterion	9	10
Not Criterion	8	20

Table 12
Contingency Table: Criterion on CLR Versus Criterion on TLR

Conservation of Length Relations (CLR)	Transitivity of Length Relations (TLR)	
	Criterion	Not Criterion
Criterion	9	15
Not Criterion	9	14

Examination of Table 11 revealed that there were 8 children in the study who met criterion for transitivity but not conservation of matching relations. From Table 12 it may be observed that 9 children met criterion for transitivity but not for conservation of length relations. In each case, about one-half of the children who could use the transitive property within a relational category failed to conserve the relations of the same category. Thus, no evidence is provided by these data that, for the children in this study, the ability to conserve relations precedes the ability to use the transitive property. The case is different, however, in the case of the Transitivity Problem.

For consideration of whether the ability to conserve matching relations precedes the ability to conserve length relations, frequencies of children who achieved the criteria for CLR and CMR are presented in Table 13. Table 14 contains frequencies with which children in the study met the criteria for TLR and TMR as an indication of whether matching precedes length in development of the transitive property of relations. The calculated chi-squares were 6.33 for Table 13 and .87 for Table 14, each with 1 *df*. The value for conservation was significant ($p < .05$), and thus a relationship between meeting criterion on CMR and meeting criterion on CLR is indicated. These results are consistent with the significant correlation of .41 between CMR and CLR and the nonsignificant correlation between TMR and TLR.

Table 13
Contingency Table: Criterion on CLR Versus Criterion on CMR

Conservation of Length Relations (CLR)	Conservation of Matching Relations (CMR)	
	Criterion	Not Criterion
Criterion	13	10
Not Criterion	5	19

Table 14
Contingency Table: Criterion on TLR Versus Criterion on TMR

Transitivity of Length Relations (TLR)	Transitivity of Matching Relations (TMR)	
	Criterion	Not Criterion
Criterion	8	10
Not Criterion	9	20

The data of Table 13 gave no indication that conservation of matching relations precedes conservation of length relations for the children in this study. In fact, 10 children who met the criterion on CLR failed to achieve the criterion on CMR. On the other hand, there were 5 children who met criterion on CMR but failed to meet criterion on CLR. This evidence is in opposition to the suggestion that the ability to conserve matching relations precedes the ability to conserve length relations.

From Table 14 it may be observed that 9 children used (as defined by the criterion) the transitive property of matching relations but not length relations. On the other hand, 10 children used the transitive property of length relations but not of matching relations. These data gave no indication that, for the subjects of this study, the ability to use the transitive property in one relational category consistently preceded the ability to use the transitive property in the other relational category.

Presumably, a solution of the Transitivity Problem required use of the transitive property of the relation "as many as." However, other abilities were necessary for a solution. Thus, the fact that some children achieved the criterion on TMR but did not reach a solution in the Transitivity Problem is consistent with the logical conclusion.

What appears inconsistent with the logical conclusion is that seven children solved the Transitivity Problem but failed to reach the criterion on the transitivity (TMR) test (see Table 8). Of these seven, however, four made a score of three on the TMR test and thus gave evidence of some facility in transitivity. The failure of the other three children may be attributed to inaccuracy of measurement.

Another discrepancy between the data and the logical conclusion is the fact that 8 children used the transitive property (as defined by the criterion on TMR), but did not conserve matching relations. It is interesting to note that 5 of these 8 children were in the full treatment group. It is also of interest to observe that in the entire study, 13 children who had full treatment achieved the criterion on TMR while only 4 children in the partial treatment did so.

Discussion and Conclusions

The Effectiveness of the Treatment

The mean performance of the children in the full treatment group was significantly greater than the mean performance of the children in the partial treatment group on the Transitivity of Matching Relations Test. This was an indication that the treatment was effective in improving the ability of the children in using the transitive property of these relations. However, the results from the Transitivity Problem indicated no relationship between a student's membership in a treatment group and his level of performance on the Transitivity Problem. This apparent discrepancy may be interpreted by an examination of the tasks and the instructional activities. In the instructional setting the children were instructed to establish the relation between two sets, say A and B, and between B and a third set, C. The sets were constructed in such a way that the same relation existed between B and C as between A and B. The children were then asked to predict the relation between A and C and were given an opportunity to verify their prediction. Each item of the structured transitivity test followed this same procedure except that on the test the child did not have the opportunity to verify his conclusion. Also, in the testing situation the objects were screened at the time of the transitive inference, whereas this was not always the case in instruction. In the Transitivity Problem, the child was required to compare sets A and B, and sets A and C where A contained two more objects than B or C. He then was required to remove (either physically or mentally) two objects from the set A to form a new set which was equivalent to B and C before applying the transitive property of "as many as," and to conclude that B was equivalent to C. The reasonable conclusion then, is that the treatment improved the ability of the children to perform tasks very much like the treatment activities, but this improvement did not generalize to the Transitivity Problem, a higher order task.

These results are consistent with previous transitivity training studies. In a study with five- to seven-year-old children, Smedslund (1963c) found that none of the children acquired transitivity of weight due to practice. In another study he (Smedslund, 1963c) found that about 30% of a group of eight-year-old children acquired transitivity of weight by practice, while only 12.5% of a control group acquired transitivity. Thus, behavior indicative of transitivity has been obtained in some training studies, but it appears to be difficult to induce transitivity by practice.

It appears from Piaget's theory that if a child's cognitive structure contains the grouping of addition of asymmetrical, transitive relations, he can use the transitive property of any such relations, regardless of the concrete embodiment. Piaget (1952, p. 204) has indicated, on the contrary, that a formal structure of transitivity is not acquired all at once, but it must be reacquired every time a new embodiment is encountered. Sinclair (1971) has further suggested that properties of the concrete embodiments (such as discrete or continuous) will affect the attainment of psychologically parallel concepts.

In the present study, experiences in length relations were given to introduce an embodiment of the transitive relations in addition to the matching relations, but no instruction was given in transitivity of the length relations. The results indicate that while the treatment improved the ability to use transitivity of the matching relations, there was no corresponding improvement in the ability for the children to use transitivity of length relations. Thus, the conclusion was reached that the treatment was rather task specific and no generalized scheme of transitivity was induced.

This conclusion is consistent with Piaget's conjecture, and with the results of training studies in conservation. For example, Beilin's (1965) subjects improved in conservation of number and length when experiences were given. However, the training was not sufficient to foster generalization to conservation of area.

The results of the Asymmetric Property of the Matching Relations Test indicate that the treatment was effective in improving the ability of the children in the full treatment group in using the asymmetric property of the matching (order) relations. This may be interpreted not as a transfer of training, but as a direct consequence of the instructional activities. In each activity, the instructors stressed the relations which did not hold as well as the relation which did hold. Consider, for example, an activity in the differential treatment in which there were more a's than b's. After the transitive inference or conservation question, "Are there more *a*'s than *b*'s?" the instructor also asked "Are there as many *a*'s as *b*'s?" and "Are there fewer *a*'s than *b*'s?" If a child failed to answer "no" to each of these latter two questions, the instructor corrected the child by using the materials. The statement that there are not fewer *a*'s than *b*'s is equivalent to the statement that there are not more *b*'s than *a*'s. This logical equivalent (that there are not more *b*'s than *a*'s) is precisely the asymmetrical inference from the relation which does hold: there are more *a*'s than *b*'s. This situation may have been interpreted in this way by the children, so that the treatment effect was obtained for the asymmetric property.

The differential treatment contained four lessons on conservation of matching relations and five lessons on transitivity of matching relations. The conservation portion of the treatment was not successful in improving the conservation ability of the children in the treatment group. Many of the conservation training studies previously reported have indicated that conservation ability has been improved (Beilin, 1971). The conservation treatment in the present study was apparently either too short, or the activities were inappropriate for the subjects of the study. Another possible factor was that the transitivity instruction intervened between the conservation instruction and the testing period. This delayed the testing on conservation for one more week after instruction than the testing on transitivity. There remains the possibility that the conservation lessons were instrumental in fostering the improvement of performance of the treatment group in the transitive and asymmetric properties.

Matching and Length Relational Properties

The mean performance of the six-year-old group was higher than the mean performance of the five-year-old group on all matching relations tests except transitivity. It is not surprising that these cognitive abilities improved between the ages of five and six. The amazing result is that age had no significant effect on the abilities of the children in using any of the length relational properties. Consideration of the means indicated that performance on length relational properties was at about the same level as performance on matching relational properties. Consideration of criteria levels showed that more children attained conservation and transitivity of length relations than the corresponding properties of matching relations. Thus, from the point of view of relations rather than number and length, Sinclair's (1971) hypothesis is not confirmed for the children in this study.

Conservation and Transitivity Attainment

The result that about one-half of the children who used the transitive property in each relational category failed to use conservation of that respective category is at variance with results of previous studies. Smedslund (1964) found only 4 of 160 subjects who passed the test on transitivity and failed on conservation of discontinuous quantities, and only 1 subject was in the corresponding cell for length. Owens and Steffe (1972) observed only 4 of 126 instances (among 42 subjects) in which transitivity of a matching relation preceded conservation of that relation. Divers (1970) found that in 87% of the cases where transitivity of a length relation was attained, the relation was also conserved. In the studies cited, the results consistently indicated that attainment of conservation preceded attainment of the transitivity property. None of the studies involved instruction or practice, and the present results may be interpreted in terms of the treatment effect. The treatment was effective in improving performance on the test of the transitive property while the treatment had no effect on conservation performance for matching relations. Thus, some children in the treatment group met the criterion on the transitivity test who might otherwise not have attained transitivity. Only two children who used transitivity on the Transitivity Problem failed to exhibit conservation. This explanation applies, however, only to the matching relational category, because the treatment was not effective in improving the performance on transitivity of length relations.

Perhaps an interpretation can be made in terms of the characteristics of the children in the sample. Skypek (1966) conducted a study which involved both middle and lower socio-economic status children. It was found that among the low status children, the development pattern of cardinal number conservation was erratic. While the present study included no middle class group for comparison, it appears that the pattern of attainment of conservation and relational properties was irregular for these low economic status subjects.

MARTIN L. JOHNSON

Learning of Classification and Seriation by Young Children

The acts of classifying and ordering objects may be analyzed both psychologically and mathematically. Beth and Piaget (1966) have attempted to explain these acts psychologically by postulated models of cognition called "groupings." Two of the Groupings, Grouping I and Grouping V, deal with classes and asymmetric relations, respectively. These two groupings provide models for the cognitive acts of combining individuals in classes and assembling the asymmetrical relations which express differences in the individuals, or more specifically, models for classification and seriation.

The elements of Grouping I are classes which are hierarchically arranged. Somewhere between late preoperational and early or middle concrete operational stages the child can readily ascend such a hierarchy of classes by successively combining elementary classes into supraordinate classes ($A + A' = B$; $B + B' = C$, etc.) (Flavell, 1963). Furthermore, the child can just as easily descend the hierarchy, beginning with a supraordinate class and decomposing it into its subordinate classes, ($C - B' = B$, etc.). In addition, the child can destroy one classification system in order to impose a new and different one on the same data (Inhelder and Piaget, 1964).

Beth and Piaget (1966) point out that seriation behavior can be found in children from the sensory-motor stage onward, with operational seriation coming only after the child begins to anticipate the series and coordinate the relations involved in forming the series. An example is where a child is given elements A, B, C, D, E, F, G, etc., and is asked to arrange them according to some asymmetrical transitive relation. If the child finds a systematic method -- puts down the "smallest" of the elements (A) after carrying out pairwise comparisons, the "smallest" (B)

The experiment reported in this chapter is based on a doctoral dissertation in the Department of Mathematics Education at the University of Georgia (Johnson, M. L., 1971).

of the remainder (again after comparing in pairs), then again the "smallest" (C) of the remainder, etc. -- he understands in advance that any element E will be at the same time "larger" than any element already put down (A, B, C, D) and "smaller" than the remaining elements (F, G, etc.). Grouping V provides a model for the cognitive actions present in such an act of seriating objects.

Mathematically, classification and seriation can be interpreted as being logically dependent on equivalence and order relations. It is well known in mathematics that each equivalence relation serves as the basis for a classification of objects (Herstein, 1964) and conversely that every exhaustive classification of objects into distinct classes defines an equivalence relation, where objects are considered as being equivalent if they are classified together. For example, given a set of linear objects (A), with the instructions to "put all objects having the same length together," the set can be partitioned into disjoint subsets by using the mathematical properties of "same length as." On the other hand, nonreflexive, asymmetrical, transitive relations can serve as a basis for ordering a collection of objects, or for the act of seriation.

Classificatory behavior of young children has been the subject of a substantial amount of research in recent years. Inhelder and Piaget (1964) were among the first to systematically study the behavior of children as they attempted to form classes. These authors report behavior related to classificatory acts ranging from "graphic collections" (Stage I) in which the child forms spatial wholes, to true classification (Stage III). True classification appears when children are able to coordinate both the intension and extension of a class as shown by an ability to solve class inclusion problems -- somewhere around 8-9 years of age. Lovell, Mitchell, and Everett (1962) found behavior similar to that found by Inhelder and Piaget with only Stage III children being able to group objects according to more than one criterion; such as color, shape, or form. The fact that the basis of classification children use is age related was revealed by Olver and Hornsby (1966). Their research showed that collections made by very young children are based on perceptible properties of objects (color, shape, etc.) with an increase of functional based equivalence as children grow older. Other researchers (Maccoby and Modiano, 1966) reported that the choice of criteria for classification is a function of the child's culture. While this may be the case, Olmsted, Parks, and Rickel (1970) reported that the classification skills of culturally deprived children, including an increase in the variety of criteria used for classification, could be improved by involving the children in a systematic training procedure. Edwards (1969) also reported an increase in classification performance of children due to training. Other investigators (Clarke, Cooper and Loudon, 1969; Darnell and Bourne, 1970) reported that conditions of training, such as making the child aware of natural relationships or orderings among a set of objects, may facilitate the learning of equivalence relations.

Seriation behavior develops in stages similar to classification behavior (Inhelder and Piaget, 1964). Systematic or operational seriation appears in Stage III -- approximately eight years of age. Operational seriation is distinguished by (1) the discovering of a systematic way of forming a series and (2) the ability to insert new elements in an existing series without relying on trial-and-error procedures. Developmental stages consistent with the findings of Inhelder and Piaget have been reported by Elkind (1964) and Lovell, Mitchell, and Everett (1962).

Very little research has been reported in which training procedures were used in an attempt to facilitate seriation ability. Coxford (1964) reported that selected instructional activities had a facilitating effect on the seriation ability of those children who were already in a transitional stage. Holowinsky (1970), however, reported that any increase in seriation ability of four-, five- and six-year-old children in his sample was more likely due to age increases than to instructional activities. Clearly, the issue of training effects has not been resolved.

The current literature indicates that various factors influence the ability of a child to determine criteria for classification. Although the training research has not been completely unfavorable, classification has been approached only as a general categorizing process not including the major action in classifying -- the forming of equivalence classes. Hence, any relationship which may exist between the child's knowledge of the mathematical properties of an equivalence relation and his classification skills based on that relation has not been explicated.

Similarly, the current literature does not explicate any relationship which may exist between seriation ability and properties of order relations. Specifically, does a relationship exist between the child's knowledge of the transitive property of an asymmetric, transitive relation and the child's ability to seriate on the basis of that relation? Furthermore, it has not been determined whether seriation ability is relational specific or material specific. Conversely, because classification and seriation involves using mathematical relations, does classification and seriation training produce an understanding of the mathematical properties of the relations involved, specifically, the transitive property?

Purpose

The main purpose of this experiment was to determine the influence of training on the ability of first and second grade children to classify and seriate objects on the basis of length. A second purpose was to investigate the influence of such training on the child's ability to conserve and use the transitive properties of the above relations "same length as," "longer than," and "shorter than."

Other objectives were to determine if the subject's ability to use the transitive property of the equivalence relation "same length as" was related to his ability to classify on the basis of the relation; to investigate the relationship between the child's ability to use the transitive property of the relations "longer than" and "shorter than" and his ability to seriate on the basis of these relations; and to determine if the ability to seriate linear objects is material specific or relational specific.

Method

The Subjects

Eighty-one subjects, comprised of thirty-nine first grade children and forty-two second grade children were chosen for this study. Twenty-three first grade and twenty-four second grade children were from the

W. H. Crogman Elementary School, while sixteen first grade and eighteen second grade children were enrolled at the Cleveland Avenue Elementary School; both schools in Atlanta, Georgia. At the beginning of this study, March 16, 1971, the mean age for first grade was 80.8 months and for second grade 91.8 months.

The W. H. Crogman Elementary School is located near downtown Atlanta in a Model Cities area. All subjects from W. H. Crogman School who participated in this study were Negro. By virtue of having been in a Model Cities area, the educational programs of W. H. Crogman Elementary were enriched by the use of parents as "teacher helpers" and a well-planned extended day program for first and second grade children. Many of the children who participated in this study took part in home economics, art, music, organized physical education, and a variety of other activities.

Cleveland Avenue Elementary School is located in Southwest Atlanta and served a predominately Caucasian, middle-class student population. Of the thirty-four children in the sample from this school, thirty-two were Caucasian and two were Negro. Within each school, the sample was randomly selected from the existing first and second grade classes.

Description of Learning Material

Two instructional units were written for this study. Unit I consisted of six lessons designed to acquaint the students with the relations "same length as," "longer than" and "shorter than" and to make proper comparisons based on these relations. Unit II consisted of ten lessons designed to give experiences in classifying on the basis of the equivalence relation "same length as" and seriating on the basis of the order relations "longer than" and "shorter than." For example, lessons 1 and 3 of Unit II were primarily concerned with classifying sticks, straws, pipecleaners, and ropes on the basis of length. In lesson 2 the child was asked to determine the longest and/or shortest object from a collection of objects given to him. The procedure followed was to make pairwise comparisons until the longest (shortest) object was determined. In subsequent lessons, the child was asked to seriate collections of sticks and then mixed collections (strings, sticks, straws, pipecleaners) using a procedure consistent with Piaget's stage three behavior. At least two lessons required the child to insert additional objects into series already formed.

Instructional Schedule and Modes of Instruction

Instruction on Unit I began at W. H. Crogman School on the morning of March 18, 1971. Similar instruction began at Cleveland Avenue School on March 19, 1971. Because of having to alternate between the schools on consecutive days, twelve days were required for the completion of Unit I. Upon completion of Unit I, a Criterion Test (see *Criterion Test*) was administered. The children who met criterion were randomly placed into either an experimental or control group. The experimental group received instruction on Unit II while the control group received no further instruction. Instruction on Unit II began on April 14, 1971 with instruction being given at W. H. Crogman School in the morning and at Cleveland Avenue School in the afternoon. Ten instructional days were required for the administration of Unit II, ending on April 27, 1971. All instruction was carried out in groups of approximately six students in 20 minute sessions. Unit II was taught only by the investigator while teacher aides helped with the instruction of Unit I.

Tests

Instruments were constructed to measure the children's knowledge of the length relations, ability to conserve and use the transitive property of the relations, ability to seriate using the order relations, and ability to classify using the equivalence relation.

Criterion Test. A nine-item test was constructed to determine if, at the end of Unit I, the children understood the relations and terms used in the conservation and transitivity tests to be administered as pretests. To meet criterion on this test, the child had to meet criterion on each of the three relations, which was defined as correctly performing on two of the three questions asked about each relation. For example, the child would be asked (from a pile of sticks with a standard stick placed before the pile) to "find a stick the same length as the standard stick," "find a stick longer than the standard stick," and "find a stick shorter than the standard stick." Similar instructions were given for the other six items which included both sticks and strings. All questions were asked in random order to each child.

Conservation of Length Relations Test (CLRT). This test consisted of six items; two each concerning the relations "same length as," "longer than," and "shorter than." Two perceptual stimuli were given for each relation; neutral and conflictive. All of the materials were red and green sticks 3/8" in diameter differing in length by 1/8" within an item. In items with the neutral stimuli a red and green stick would be displayed and the child was asked "Is the red stick the same length as the green stick?" or "Is the red stick longer than the green stick?" or "Is the red stick shorter than the green stick?" The question asked would depend on whatever relation did hold between the two sticks. After the child had determined which relations did hold, one stick was moved right or left so that the left end of one stick coincided with the right end of the other. Three questions were now asked in random order. "Is the red stick the same length as the green stick?", "Is the red stick longer than the green stick?", "Is the red stick shorter than the green stick?".

The items with conflictive stimuli were administered in a slightly different way than the items with neutral stimuli. After the child had determined the relation that existed between the red and green sticks, they were moved to form a "T" and the three questions were then asked.

To receive a score of one on an item, the child had to answer the three questions correctly. The correct sequence of answers depended on the item being given. This test was given both as a pretest and a posttest.

Transitivity of Length Relations Test (TLRT). This test consisted of six items; two each for the relations "same length as," "longer than," and "shorter than." Two perceptual stimuli were present; screened and conflictive. All materials in this test consisted of red, blue, and green sticks all 3/8" in diameter and differing in length by 1/8". In each item, the child had first to determine the relation that existed between the red and blue sticks, then the blue and green sticks. To make an inference about the relation that existed between the red and green sticks the child was again asked three questions in random order as in the CLRT. On the items with screened stimuli the final inference about the length of the red and green sticks had to be made with the sticks in boxes and not visible by the subjects. This test was used both as a pretest and a posttest with scoring as in the CLRT.

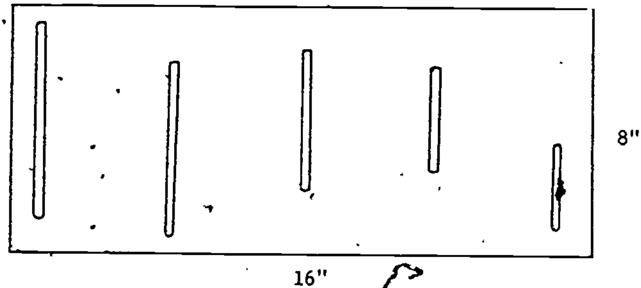
Seriation Test. A 12-item test was constructed to assess the child's ability to seriate on the basis of "longer than" and "shorter than." Items 1-6 were based on the relation "longer than;" items 7-12 were based on "shorter than." Item 1 and item 7 required the child to seriate six sticks (free seriation), all 3/8" in diameter, differing in length by 1/8" with the shortest stick being 5 1/2" long. Item 2 and item 8 required the child to seriate six strings (free seriation) of the same length as the sticks in item 1 and item 7.

For the free seriation items, a point was given for each stick or string judged to be in the "correct place" with respect to the relation given. For example, when the child had indicated that his series was formed, he was then asked to show how the objects were in order from the longest to shortest (shortest to longest). Now, if, for instance, he was basing his ordering on "longer than," and he indicated that his series was formed from left to right, a point would be given for a stick *b* if *b* was shorter than the stick it immediately succeeded and at the same time longer than the stick it immediately preceded. A maximum of four points was awarded for each of the free seriation items.

Items 3, 4, 5, 9, 10 and 11 required the child to insert a stick into a series already formed. However, the sticks in the series were glued on an 8" X 16" piece of cardboard, spaced and staggered so that a baseline was not discernable, as in Figure 1.

Figure 1

Item #4 (seriated from longest to shortest - left to right).



The series in items 3 and 9 consisted of four sticks, in items 4 and 10 five sticks, and in items 5 and 11 six sticks. In each case the sticks were equally spaced. Items 6 and 12 were also insertion items but the existing series had a visible baseline and the sticks could be moved about. One point was given for each correct answer.

Classification Test. This test consisted of 3 items; two requiring the child to group sticks on the basis of length and one in which the child had to determine the criteria used for sticks already grouped.

The materials for item 1 consisted of 12 green sticks, each 3/8" diameter with four of length 5", four of length 5 1/4" and four of length 5 1/2". One stick of each length was mounted on a piece of paper board. The three mounted sticks were pointed out to the child who was then instructed to "find all of the sticks that would go with this stick (5"), this stick (5 1/4") and this stick (5 1/2")." The nine sticks to be classified were in disorder before the child. A record of all sticks correctly and incorrectly placed was kept by the experimenter.

The materials for item 3 consisted of ten red sticks all 3/8" diameter, three of length 4", three of length 4 1/4", three of length 4 1/2", and one of length 4 3/4". The ten sticks were given to the child and he was instructed to "put all of the sticks together that belong together." A record of the child's actions was kept by the experimenter.

Item 2 required that the child determine the criteria used for grouping. The materials for this item consisted of fifteen sticks; five each at length 6", 6 1/4", and 6 1/2". The sticks were placed into three distinct piles about 15 inches apart on a table. Within a pile, sticks differed in color and diameter; with length being constant. The child was instructed to "Tell me why I have all of these sticks together in this pile (6"), in this pile (6 1/4") and in this pile (6 1/2")." If a correct answer was given, the child was asked to justify his answer. Upon justification, he was then asked, "Why do I have these sticks in different piles?" Again a justification for a correct answer was asked for. A record of all answers was kept by the experimenter.

The Experimental Design and Statistical Analysis. Two treatment conditions within two grade levels within two schools produced eight comparison groups. Table 1 is a layout of the design. S_1 and S_2 represent W. H. Crogman Elementary School and Cleveland Avenue Elementary School, respectively. The numerals "1" and "2" represent grades 1 and 2. The letters E and C represent experimental and control groups, and G_i ($i = 1, \ldots, 8$) represents the eight different groups.

Because the main purpose of the criterion test was to eliminate subjects who did not have a knowledge of the relations, no test of significance was performed on the criterion test data. It should be pointed out that all children met criterion on the Criterion Test.

A 2 X 2 X 2 factorial design utilizing analysis of variance (MUGALS)* was used to determine the effect of the two classification (School and Grade Level) and treatment variables on the seriation test. An analysis

*MUGALS (Modified University of Georgia Least Squares Analysis of Variance), Athens, Georgia, University of Georgia Computing Center, 1966.

Table 1
Outline of the Design

School	Grade Level	Treatment	Tests: Crit Test	CLRT Pre	TLRT Pre	Seri Test	Class Test	CLRT Post	TLRT Post
S_1	1	E	G_1	G_1	G_1	G_1	G_1	G_1	G_1
		C	G_2	G_2	G_2	G_2	G_2	G_2	G_2
	2	E	G_3	G_3	G_3	G_3	G_3	G_3	G_3
		C	G_4	G_4	G_4	G_4	G_4	G_4	G_4
S_2	1	E	G_5	G_5	G_5	G_5	G_5	G_5	G_5
		C	G_6	G_6	G_6	G_6	G_6	G_6	G_6
	2	E	G_7	G_7	G_7	G_7	G_7	G_7	G_7
		C	G_8	G_8	G_8	G_8	G_8	G_8	G_8

of covariance (MUGALS) was used to analyze the conservation and transitivity posttest scores using the respective pretests as covariates. An item by item analysis involving the treatment variable was performed on the classification test data using contingency tables and Chi-square test of independence. To determine relationships between transitivity, seriation, and classification, a series of contingency tables was constructed and tested with Chi-square tests of independence.

Results

Seriation Test

The overall mean for the seriation test was 12.51 with a standard deviation of 7.03. The total possible score by an individual on this test was 24. Table 2 contains the means for grades and treatment within groups.

Table 2
Means for Seriation Test

School	Grade	Experimentals	Controls
Crogman	1	12.83	6.55
	2	15.50	12.42
Cleveland	1	11.25	7.88
	2	20.00	13.56

Table 3 contains the difficulties of all dichotomous items on the seriation test. Because items 1, 2, 7, and 8 were nondichotomous the difficulties of these items are reported as p-values in Table 4.

Table 3
Item Difficulties of Dichotomous Items - Seriation Test

Item No.	Frequency of Correct Response	Difficulty
3	46	.57
4	33	.41
5	33	.41
6	46	.57
9	41	.51
10	36	.44
11	43	.53
12	44	.54

Table 4
P-Values for Nondichotomous Items*

		P-Values				
Item No.	Score	0	1	2	3	4
1		.26	.06	.17	.00	.51
2		.28	.16	.23	.03	.30
7		.27	.07	.12	.05	.48
8		.30	.12	.27	.05	.26

*p is the ratio of Ss who received score i (i = 0, 1, 2, 3, 4) for item x (x = 1, 2, 7, 8) to the total number of subjects answering.

Inspection of Table 3 shows that less than fifty percent of the children were able to correctly answer items 4, 5 and 10. Items 4 and 10 involved inserting a stick into a fixed five-stick series in which the sticks were ordered from longest to shortest (item 4) and from shortest to longest (item 10). It was expected that item 5 would be more difficult than item 4, because item 5 contained six sticks as opposed to five in item 4 and were arranged in a staggered series from longest to shortest. Because items 5 and 11 were identical except for order, the difference in difficulty was not expected. Furthermore, inserting a stick into a six-stick series with a baseline (items 6 and 12) appeared to be slightly easier than inserting into a six-stick series without a baseline (items 5 and 11).

Table 4 clearly shows that more children were able to correctly seriate sticks (items 1 and 7) than strings (items 2 and 8). Little difference in difficulty was found between performance on the "longer than" item using sticks (item 1) and the "shorter than" item using sticks (item 7). Similarly, little difference in difficulty was found between performance on the "longer than" item using strings (item 2) and the "shorter than" item using strings (item 8). It also appeared that children either could not put any objects in correct order (score of 0); could correctly order up to four or make two pairs of three, each correctly ordered (score of 2); or could correctly order all six objects (score of 4).

The analysis of variance for the seriation test is reported in Table 5. Both grade ($p < .01$) and treatment ($p < .01$) were highly significant main effects. No differences could be detected due to school. No significant first or second order interactions could be detected. It is clear that the experiences provided to the subjects in this study were sufficient to improve their seriation ability as measured by the seriation test.

Table 5
Analysis of Variance for Seriation Test Scores

Source of Variation	df	MS	F
S (School)	1	35.66	<1.00
G (Grade)	1	648.79	17.22**
T (Treatment)	1	452.92	12.02*
S x G	1	42.70	1.13
S x T	1	.25	<1.00
G x T	1	.02	<1.00
S x G x T	1	48.41	1.29
Error	73	37.68	

**(p < .0005), * (p < .001)

Conservation of Length Relations Test (Posttest)

Table 6 contains the means for all groups on the CLRT posttest. An interesting observation is that the means for first-grade controls were somewhat higher than means for first-grade experimentals across schools.

Table 6
Means for Conservation of Length Relations Test (Posttest)

School	Grade	Experimentals	Controls
Crogman	1	2.50	3.20
	2	4.50	4.08
Cleveland	1	2.25	4.25
	2	5.22	4.89

The CLRT was given both as a pretest and a posttest. A comparison of item difficulties on the pretest and posttest is given in Table 7. Overall, the items on the CLRT were easier on the posttest than on the pretest. Item 1, involving "same length as," was more difficult than item 2, also involving "same length as," on both administrations of the test. This is surprising since item 1 used a "neutral" situation while item 2 was a "conflictive" item. Divers (1970) found that different perceptual situations had little effect on conservation ability. Items 3 and 4, involving "longer than," were the easiest items on the posttest.

The results of the CLRT posttest were analyzed by analysis of covariance using the CLRT pretest as a covariate. The results of this analysis are reported in Table 8.

Table 7
Item Difficulties of Conservation of Length Relations Test: Pretest and Posttest

Item No.	Relation and Situation	Difficulty Pretest	Difficulty Posttest
1	Same Length As (N)	.35	.56
2	Same Length As (C)	.42	.63
3	Shorter Than (N)	.63	.67
4	Shorter Than (C)	.43	.70
5	Longer Than (N)	.35	.64
6	Longer Than (C)	.41	.64

Table 8
Analysis of Covariance for Conservation of Length Relations Test Scores (Posttest)

Source of Variation	df	MS	F
S (School)	1	14.45	6.39*
G (Grade)	1	10.68	4.72*
T (Treatment)	1	.67	<1.00
S x G	1	8.98	3.97
S x T	1	.42	<1.00
G x T	1	.34	<1.00
T x S x G	1	.98	<1.00
Error	72	2.26	

*(p < .05)

The main effects of school and grade were both significant. In view of past research on conservation ability, it was expected that older children would be better conservers of length relations than the younger children; however, it was not expected that the school effect would be significant. No significance could be detected due to treatment. No statistically significant interactions were found; however, there was a possible suggested interaction between school and grade.

Transitivity of Length Relations Test (Posttest)

Table 9 contains the means for all children on the TLRT posttest. Item difficulties for the pretest and posttest are given in Table 10.

Inspection of Table 10 reveals that on the pretest, all items except item 3 were of near equal difficulty. It was not expected that item 3 would be easier than item 4 because item 3 required the child to make an inference about the relative length of sticks placed in boxes and not visible to the child. An interesting result was the change in difficulty of items 1 and 2 in a positive direction from pre-to posttest and a change in difficulty in a negative direction for items 3, 4, 5, 6 from pretest to posttest. The items involving the linear order relations were at least as difficult after the extensive training on strategies of seriation utilizing these relations.

Table 9
Means for Transitivity of Length Relations Test (Posttest)

School	Grade	Experimentals	Controls
Crogman	1	1.50	1.40
	2	2.50	1.33
Cleveland	1	2.13	2.50
	2	2.67	3.44

Table 10
Item Difficulties of Transitivity of Length Relations Test: Pretest and Posttest

		Difficulty	
Item No.	Relation and Situation	Pretest	Posttest
1	Same Length As (S)*	.33	.49
2	Same Length As (C)	.30	.43
3	Shorter Than (S)	.44	.38
4	Shorter Than (C)	.31	.28
5	Longer Than (S)	.33	.22
6	Longer Than (C)	.31	.30

*S: Screened, C: Conflictive

Table 11
Analysis of Covariance for Transitivity of Length Relations Test Scores (Posttest)

Source of Variation	df	MS	F
S (School)	1	14.96	6.31*
G (Grade)	1	2.12	<1.00
T (Treatment)	1	.59	<1.00
S x G	1	.001	<1.00
S x T	1	8.14	3.43
G x T	1	.12	<1.00
T x S x G	1	1.40	<1.00
Error	72	2.37	

*(p < .05)

Table 11 contains the results of the analysis of covariance on the TLRT posttest with the TLRT pretest used as a covariate. Only the main effect of school was significant. No significant interactions were detected with only a possible interaction suggested between school and treatment.

The test statistics for the seriation test, the CLRT pre- and posttest, and TLRT pre- and posttest are contained in Table 12. Test correlations are given in Table 13. An unexpected result was that the seriation

Table 12
Test Statistics for Seriation, Conservation, and Transitivity Tests (N = 81)

Test	Number of Items	Possible Score	Mean	SD	Reliability
Seriation	12	24	12.5	7.03	.81**
CLRT (pretest)	6	6	2.58	1.98	.77*
CLRT (posttest)	6	6	3.84	1.98	.85*
TLRT (pretest)	6	6	2.03	1.47	.46*
TLRT (posttest)	6	6	2.11	1.66	.63*

**Alpha Coefficient, *KR-20

Table 13
Test Correlations

	Con. Pre.	Con. Post	Tran. Pre.	Tran. Post
Seriation	.21	.25*	.15	.26*
Con. Pre.		.65*	.45*	.38*
Con. Post			.43*	.33*
Tran. Pre.				.31*

*($p < .01$)

test did not correlate more than .26 with any other test. All correlations differed significantly from the zero correlation except the correlations between the seriation test and the conservation and transitivity pretest.

Classification Test

Item 1 on the classification test required the child to find and group into three distinct piles sticks similar to a given stick. From the children's responses, four performance categories were identified. They were: (a) the child did not attempt to classify sticks; (b) the child made some partial classes but did not exhaust the set of sticks to be classified; (c) the child exhausted the set but made some incorrect choices; and (d) the child correctly classified all sticks. Table 14 shows the number of subjects exhibiting each of the above four types of performance on item 1 by treatment, grade, and school. A Chi-square test of independence was performed for each main effect and performance. The control subjects performed comparably to experimental subjects on item 1. Thirty-six of forty-seven (76%) subjects at Crogman School were able to classify all sticks. The same percent of category d responses was found at Cleveland School; twenty-six of thirty-four (76%). A slight relationship ($\chi^2 = 6.78$, $p < .10$) was found between performance on item 1 and grade level.

Item 2 required the subjects to discover the criteria for classification. In this item sticks of different colors and diameters were presented in three distinct piles and the child was asked to give a reason for their being grouped together in separate piles. The child was also asked to tell why distinct piles were formed. Five distinct categories of

Table 14
Frequency Table: Performance on Item 1 Contrasted
With Treatment, Grade, and School

		Performance			
		a	b	c	d
Treatment	E	1	7	4	29
	C	0	4	3	33
Grade	1	1	8	5	25
	2	0	3	2	37
School	Cr	1	7	3	36
	Cl	0	4	4	26

responses were identified. They were: (a) the child did not discover the criteria; (b) the child gave a correct reason for the piles being together but without justification; (c) a correct reason was given with justification; (d) in addition to justifying the reason for sticks belonging in distinct groups, the subject correctly gave a reason for sticks being in different groups but without justification for his reason; (e) all of (d) with justification. The overall performance of the subjects on item 2 is presented in Table 15 contrasted by treatment, grade, and school. A slight relationship was found ($\chi^2 = 9.26$, $p < .10$) between performance and treatment; however, a higher frequency of category (e) responses was given by the control subjects with a reversal for category (d) responses. Overall, it can be seen that about 75% of the subjects failed to discover the criteria for classification in item 2.

Table 15
Frequency Table: Performance on Item 2 Contrasted
With Treatment, Grade, and School

		Performance				
		a	b	c	d	e
Treatment	E	29	1	2	6	3
	C	30	0	2	0	8
Grade	1	29	0	3	1	6
	2	30	1	1	5	5
School	Cr	35	1	3	2	6
	Cl	24	0	1	4	5

In item 3 the child was given ten sticks and asked to classify them as he desired where the sticks differed only in length. One stick was longer than all of the others, requiring the child to come to grips with forming a class with one element. Four categories of performance were identified: (a) no attempt was made to group the sticks; (b) the child made at least two piles with the sticks being placed incorrectly; (c) the child put all sticks in correct piles according to length except the longest sticks; (d) the child correctly classified all sticks, including the longest stick. The overall performance of the subjects contrasted by treatment, grade, and school is given in Table 16.

Table 16
Frequency Table: Performance on Item 3 Contrasted
With Treatment, Grade, and School

		Performance			
		a	b	c	d
Treatment	E	5	17	3	16
	C	1	21	4	14
Grade	1	2	22	5	10
	2	4	16	2	20
School	Cr	4	22	6	15
	Cl	2	16	1	15

No significant relationship could be detected between performance on item 3 and treatment, grade, or school. However, the frequencies reported for grade 2 shows that second graders gave more correct category (d) responses than the first graders which indicates that more older children were able to deal with a class consisting of one member than their younger counterparts.

Classification and Transitivity

One purpose of the study was to investigate relationships between the ability to use the transitive property of "same length as" and classification ability on the basis of length. Tables 17, 18 and 19 contain the subjects' classification responses on items 1, 2 and 3, respectively, partitioned by transitivity score. The transitivity score was a result of the subject's performance on the transitivity items involving "same length as" on the TLRT posttest. Zero, one, and two were assigned as transitivity scores depending upon whether the subject correctly answered none, one, or both of the transitivity items. In order to increase cell frequencies, rows indicating intermediate levels of performance on the classification test were combined as explained by Guilford (1956). Tables 17 and 18 show that performance on item 1 and 2 was slightly related to

Table 17
Contingency Table: Classification Performance (Item 1) vs
Transitivity Ability (same length as - Posttest)

Classification Performance	Transitivity Score		
	2	1	0
d	25	16	21
a–c	3	3	13

$\chi^2 = 5.73$, $p < .10$

transitivity ability of "same length as." No relationship could be detected between transitivity ability and classification performance on item 3. Perhaps transitivity was not needed to correctly perform the items on the classification test. However, the data in Table 17 indicate that at least 80% of the time, the child who scored 1 or 2 on the transitivity test performed at the highest level on the classification item.

In contrast, the data in Table 18 suggest that at least 85% of the time, the child who scored 1 or 0 on the transitivity test performed at the lowest level on classification item 2.

Table 18
Contingency Table: Classification Performance (Item 2) vs Transitivity Ability (same length as - Posttest)

Classification Performance	Transitivity Score 2	1	0
d-e	10	2	5
a-c	18	17	29

$\chi^2 = 5.72$, $p < .10$

Table 19
Contingency Table: Classification Performance (Item 3) vs Transitivity Ability (same length as - Posttest)

Classification Performance	Transitivity Score 2	1	0
d	13	7	10
a-c	15	12	24

$\chi^2 = 1.91$, $p < .50$

Seriation and Transitivity

The relationship between seriation ability using the relations "longer than" and "shorter than" and the ability to use the transitive properties of these relations was also investigated. These results are presented in Tables 20 and 21. The transitivity score refers to whether

Table 20
Contingency Table: Seriation (longer than) vs Transitivity (longer than - Posttest)

Seriation Score	Transitivity Score 2	1	0
9-12	4	12	12
5-8	1	6	13
0-4	2	10	21

$\chi^2 = 3.91$, $p < .50$

Table 21
Contingency Table: Seriation (shorter than) vs
Transitivity (shorter than - Posttest)

Seriation Score	Transitivity Score		
	2	1	0
9-12	2	10	12
5-8	5	11	14
0-4	7	5	15

$\chi^2 = 4.96$, $p < .30$

the child correctly answered none, one or both of the items on the transitivity test involving the relation "shorter than" or "longer than." The seriation score for each order relation ranged from zero to twelve.

The results of Chi-square tests of independence indicate that the hypothesis of independence between seriation ability and transitivity ability cannot be rejected beyond the .10 level of significance.

Other Relationships

It was expected that the ability to seriate sticks was related to the ability to seriate strings across relations. The results reported in Tables 22 and 23 show that a high relationship does exist between these two abilities. Inspection of Table 22 shows that of the 24 subjects who received a score of four, representing the correct seriating of six strings from shortest to longest, twenty-three also received a score of four for

Table 22
Contingency Table: Seriation of Sticks vs Seriation
of Strings (Shorter Than)

Strings	Sticks		
	4	3-2-1	0
4	23	0	1
3-2-1	17	10	7
0	1	9	13

$\chi^2 = 41.03$, $p < .001$

seriating sticks from shortest to longest. However, of the forty-one receiving a score of four for seriating sticks from shortest to longest, only twenty-three received four for seriating strings. Similarly, Table 23 reveals that of 21 students receiving a score of four with strings, 18 received a score of four for sticks while of 39 who received a score of four for sticks, only 18 received a score of four for strings, all seriation based on the relation "longer than." Not only were the abilities related, but clearly seriating strings was somewhat more difficult than seriating sticks.

Table 23
Contingency Table: Seriation of Sticks vs Seriation of Strings (Longer Than)

	Sticks		
Strings	4	3-2-1	0
4	18	3	0
3-2-1	18	15	3
0	3	2	19

$\chi^2 = 54.54, p < .001$

Table 24
Contingency Table: Insertion Ability (Baseline vs Non-Baseline)

	Non-Baseline		
Baseline	2	1	0
2	17	10	8
1	4	9	7
0	2	11	13

$\chi^2 = 13.83, p < .01$

A strong relationship was also found between the abilities to insert a stick into an existing series of sticks with a baseline and without a baseline, shown in Table 24. Seventy-four percent of the subjects who correctly inserted the stick into the two non-baseline items also correctly inserted the stick into the two baseline items while only fifty-four percent of the subjects who inserted correctly into the two baseline items could also insert correctly into the two non-baseline items.

Discussion

The results of this study clearly confirm the hypothesis that seria-ability of "linear" objects can be improved by training. It is also clear that seriation ability improves with age and, if trends hold, little ability to seriate "linear" objects can be expected below six years of age. The experiences provided in this study to the first grade children were sufficient to cause their mean performance (12.04) on the seriation test to be comparable to the mean performance of the second grade children who did not have the experiences (13.49). Being black or white appeared to have little or no effect on the subject's seriation ability.

The extent of th[illegible]cts' seriation ability, in terms of being operational in a Piagetian sense, must be questioned when one considers the overall performance on the transitivity test. In particular, the treatment appears to have had no effect on the children's ability to use the transitive property of the order relations involved in the study. In fact, no significant relationship could be detected between transitivity of "longer than" and "shorter than" and the ability to seriate using these relations. This finding is not consistent with the hypothesis presented by Beth and Piaget (1966) and confirmed by Elkind (1964) that transitivity is necessarily present when a child exhibits behavior characterized as stage three seriation behavior. The question is raised concerning what is "operational" seriation behavior. In this study, children were able to seriate strings and sticks, as well as insert additional sticks into a series already formed without any trouble but could not use the transitive property of "longer than." Such responses would indicate that the seriation training was successful in training the children to use an algorithm which was not part of an operational scheme. If this was the case, it would be expected that the relationship between seriation and transitivity would be negligible. If, however, the children were now "operational" then these findings suggest that contrary to Piaget's hypothesis, seriation behavior does not necessarily imply transitivity. In any case, it is clear that we need additional guidelines as to what constitutes operational behavior and more effective ways of measuring such behavior.

Throughout the training sessions outlined, it was frequently pointed out that if object *a* was the same length as object *b*, then the spatial position of *a* and *b* did not alter this relationship. Furthermore, if *a* was longer (or shorter) than *b*, then *a* would remain longer (or shorter) than *b* regardless of their spatial position. Even though such procedures were part of the classification and seriation training little or no difference was detected between the experimental and control groups in the performance on the CLRT. However, a significant school and grade effect was found. While the school and grade differences were expected in view of past studies, the non-significant treatment effect was unexpected. Although the procedures in this study differ somewhat from the procedures used in a study by Sigel, Roeper and Hooper (1966), they report that classification training improved ability to conserve quantity. Carey and Steffe (1968) report that selected experiences significantly improved the ability of four- and five-year-old children to conserve length. The experiences provided by Carey and Steffe were similar to the experiences provided to the sample in this study.

The results of the classification test indicate that it was somewhat easier for children to classify sticks on the basis of self-selected criteria than to discover the criteria used for sticks already classified. While little difference was found in performance (as noted by frequencies of response) on items one and three due to school and treatment, it was clear that second grade children did better on both of the items. On item three, the difference in response frequencies indicated that second grade children were able to form a class with only one element more consistently than the first graders. This finding was consistent with Piaget's observation that the concept of a singular class appears in a child around eight or nine years of age.

The hypothesis of a relationship between the child's classification ability and his ability to use the transitive property of the equivalence relation of "same length as" was not confirmed. The lack of a relationship may be explained, at least partially, in two ways: (1) A two-item test may not give a true assessment of transitivity ability. Past research reveals that much controversy exists over methodological issues and at the age at which children acquire the transitive property. Braine (1959), using a non-verbal technique, reported that children can use the transitive property of length relations as early as four and one-half years of age. On the other hand, Smedslund (1963b) reports that operational transitivity occurs around seven years of age and that Braine failed to assess transitivity. (2) Transitivity was not needed to do the classification tasks. In the case of item one, this could possibly have been the case since over half of the subjects receiving a score of zero on the transitivity test (indicating failure to correctly answer both transitivity items) performed at the highest level on this item. On item 2, over 50% of the subjects performed at the lowest level of performance across transitivity scores. Over half of the subjects receiving zero on transitivity also performed at the lowest levels of performance on item 3. Such results suggest that transitivity was not necessary for the classification items in this test.

In view of the findings of this study, questions may be raised concerning the feasibility of placing certain topics and activities in the early elementary mathematics curriculum. Consider the topic of formal linear measurement which is now being introduced by some curriculum developers as early as first grade. According to Piaget, Inhelder, and Szeminska (1960), prerequisite to understanding linear measurement is the ability to conserve length and to use the transitive property of length relations. The findings of this study indicate that about half of the first and second grade children used did not show evidence of these prerequisites. In view of this, perhaps, as pointed out by Huntington (1970), the teaching of formal linear measurement should be delayed until approximately third grade.

The idea of ordering numbers (such as 3 comes after 2 and 3 comes before 6) is one commonly taught at first and second grade. It seems reasonable that children at these grade levels would also have many experiences in ordering sets of physical objects. Certainly, such an activity is less abstract than ordering cardinal numbers per se. An example of such an ordering would be to order sticks on the basis of "longer than." This study has shown that many children at first and second grade cannot perform such ordering, causing one to question whether the child has a concept of "five," "six," "seven," etc. when he arranges them in order or if he is just recalling the order from his rote counting process. It has been shown that seriation ability, as related to linear objects, can be improved with certain experiences. It still needs to be shown whether similar results can be found with other relations.

The early elementary mathematics curriculum includes activities in forming and describing "sets" and operations with "sets." However, basic to forming sets of objects is the notion of classifying objects on the basis of certain properties of qualitative characteristics of the objects. As noted, Inhelder and Piaget (1964) have shown that children go through various stages in determining criteria for grouping. This study has shown that while children of six or seven years of age can sort sticks on the basis of length, they experience great difficulty when given a collection of sticks already partitioned on the basis of length and asked to tell why they were grouped together. This finding implies that children will experience difficulty in determining the reason or reasons X number of objects has been placed in a set. For example, suppose A = {January, February, March, April, May, ___________}. Will the child "discover" the criteria for grouping and add June to the set? Teachers should be aware that this type of problem may be quite difficult for six- and seven-year-old children.

Finally, the present study has raised questions concerning relationships between classification, seriation and transitivity ability. While some answers are given, it is not at all clear what kind of experiences children should have between the ages of four and eight in order to facilitate development of structures needed in logical activities.

RICHARD A. LESH

The Generalization of Piagetian Operations as It Relates to the Hypothesized Functional Interdependence between Classification, Seriation, and Number Concepts

Piaget's analysis of the cognitive evolution of number and other logical-mathematical concepts relies heavily on the psychological visbility of analyzing, ordering, and equating concepts (or tasks) on the basis of their underlying operational structures. Two studies, a pilot study and a training study, were conducted in order to investigate the legitimacy of Piaget's emphasis on the operational nature of mathematical concepts. In the pilot study, a Piagetian task analysis was used in order to obtain three parallel sequences of tasks which were graded in difficulty and which pertain to seriation, number, and classification concepts, respectively. These three sequences of tasks were used to investigate the interdependent development of classification, seriation, and number concepts. The results of the pilot study were then used to organize and interpret a training study. In the training study, the possibility of inducing learning which would transfer between logical-mathematical tasks that are characterized by isomorphic operational structures was investigated. Specifically, in the training study, an attempt was made to obtain transfer from classification and seriation tasks to number tasks and to delimit the nature of the transfer that occurred.

Before reporting the results of these studies, an interpretation of Piaget's description of the development of classification, seriation and number concepts is given. In this interpretation, the operational nature of logical-mathematical concepts is discussed and tasks that were used in the two empirical studies are presented.

The experiment reported in this chapter is based on a doctoral dissertation in the College of Education at Indiana University (Lesh, 1971).

Operational Concepts

One useful definition of the word "concept" can be stated as follows. A concept has been attained when one can, within a given universe of experience, distinguish instances from noninstances of the concept. On the basis of this definition, at least two subcategories can be distinguished within the class of concepts. An example of the first of these types is the concept of "red." This type of concept may be referred to as a *concrete concept* since all of the information that is necessary in order to distinguish instances from noninstances is directly given in the perceptual field. Another type of concept may be referred to as an *operational concept* in that it involves abstractions, not just from directly perceived properties of objects, but also from relations between objects, or from operations (or transformations) that are performed on objects (Piaget, 1971, p. 26).

Mathematical concepts can involve operations in at least two ways. Some concepts (e.g., set union "+") inherently involve the mastery of a system of operations. Other concepts are operational because they arise only after a system of operations has been mastered. As an example, consider the concept of a class.

The Concept of a Class

A class of objects does not exist in isolation. In order to form a class of objects C, one must be able to determine not only what elements are in C but also what elements are not in C (call this class C') relative to some subsuming class S.

A kindergarten child can be presented with three clear plastic boxes containing 8 yellow balls, 3 yellow cubes, and 8 green cubes respectively. If he is asked, "Are there more yellow things or more balls?" the response is often "more balls." An analysis of children's responses (Inhelder & Piaget, 1964) reveals that the difficulty is not that young children misunderstand the intent of the question. The difficulty seems to be that when the child's attention is drawn to the class C = (balls), the subsuming class S = (yellow things) is cognitively destroyed. Hence, the child may end up by comparing the size of the class C with the size of the class C' = (yellow cubes).

Similarly, other tasks indicate that when attention is directed toward a subsuming class, its subclasses are often confused with overlappings (Vygotsky, 1962, pp. 56-65). Based on the careful analysis of children's responses to such tasks, Inhelder and Piaget (1964, Ch. 1-4) have concluded that a concept of a class C relative to a subsuming class S requires the coordination of the operation "+" (i.e., class union, S = C + C'), with its inverse "-" (i.e., separation of classes, C = S - C').

Three Basic Types of Logical-Mathematical Operations

A group of mathematicians, the Bourbaki group (Bourbaki, 1948), wanted to isolate a small number of "matrix structures" which would be fundamental to all of the various branches of mathematics in that no one of them could be reduced to the others and that all other mathematical structures could be derived from these by combination, differentiation, or specialization. Through regressive analysis, three basic types of structures were isolated which can be roughly characterized as follows (Grize, 1960, pp. 72-81):

1. *Algebraic structures*, the prototype of which is the group. These structures were distinguished in that their form of inverse operation was *negation*.

2. *Ordering structures*, the prototype of which is the lattice. These structures were distinguished in that their form of inverse operation was reciprocity.

3. *Topological structures*, involving the concepts of neighborhood, limit and continuity.

Just as the axiomatician can analyze mathematical structures in forms of component structures and can look for the fewest and weakest axioms that will be sufficient to account for a given structure, the developmental psychologist can look at tasks that children perform and characterize them in terms of the system of operations or relations that they involve. Piagetians (Beth & Piaget, 1966, p. 186) have isolated three basic types of cognitive operations that are roughly equivalent to three types of structures determined by the Bourbaki group. Distinguished by their form of inverse, these types of cognitive operations are:

1. Operations whose form of inverse is negation, as in the set union *operation* + that gives rise to classification concepts.

2. Operations whose form of inverse is reciprocity, as in the ordering *relation* < (less than) that gives rise to seriation concepts.

3. Geometric *transformations*.

Of course, a child may not be consciously aware of the operations and relations that are implicit in his activities. For example, when a pencil A is shorter than a pencil B, and B is shorter than a pencil C, kindergarten children may be able to conclude that A will be shorter than C. Further, they can use this fact long before they are explicitly aware of the transitive property of order relations or of the system of relations that the transitive property implies. As another example of the intuitive mastery of a concept, children commonly use perfectly correct rules of grammar long before they are explicitly aware of these rules.

Intuitive Mastery of Operational Concepts

In mathematics as in language acquisition, it may be typical for children to use rules (or systems of operations) before concious awareness is attained. The intuitive mastery of a system of operations may be somewhat analogous to the acquisition of an unconscious habit. What is at first a habitual pattern (i.e., structure) for using a system of operations to achieve some end later becomes a program in the sense that various substitutes can be inserted without disturbing the overall act. The unconscious application of a system of operations becomes more and more probable in the performance of concrete tasks, and it is in this sense that one can speak of the intuitive mastery of a given operational structure.

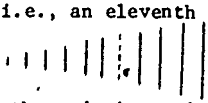
It is clear that children typically come to master a wide range of tasks that are characterized by isomorphic operational structures over a relatively short period of time. For example, in the case of the ordering relation < , a child usually becomes able to put cubes in order according to size at about the same time he becomes able to put pencils in order according to length, at about the time he can put circles in order according to diameter, etc.

A question which arises is, "How might one go about teaching children an intuitive understanding of the order relation < ?" In order to teach a concept such as "red," the child can simply be shown examples and counterexamples of red objects. However, in order to give a child an intuitive understanding of the relation < , the situation is not as simple.

The Concept of a Series

Consider the relation < as it pertains to the task of putting ten 3/8" dowel sticks (varying in length by 1 cm , the shortest of which is 9 cm) in order according to length (Inhelder & Piaget, 1964, Ch. 9).

The earliest responses that children are able to give when confronted with such a task consists of unconnected, uncoordinated pairs of shorter and longer sticks (ı| ı| ı| | ı| ı|). Later, children are able to produce two or three unconnected subseries (ı | | ı| | ı |). This is accomplished by choosing the stick that is apparently shortest (i.e., usually without making any active comparison), then some stick that is longer than the first stick selected, followed by a stick that is longer than the second one selected, etc., until the child is forced to choose a stick shorter than the last one selected. At this level of mastery of the seriation task, the child is often unable to select the shortest stick first, then the shortest of those remaining sticks, etc., until all of the sticks have been put in order. Such a response would require that the child be able to coordinate the relation "longer than" (i.e., longer than the sticks already selected) with the relation "shorter than" (i.e., shorter than those sticks remaining). Indeed, even if a child is able to correctly seriate a collection of sticks, he may still be quite unable to insert a "forgotten" stick (i.e., an eleventh intermediate stick in the series)



without breaking up the ordering and reconstructing the entire series. Insertion of a "forgotten" stick requires simultaneous consideration of the relation "longer than" and its inverse "shorter than." Thus, the concept of a series involves the gradual coordination of the relation "longer than" and its inverse "shorter than."

Mathematically, the relation that characterizes the above seriation task is a strict partial ordering, the formal definition of which follows.

Definition: < is a strict partial ordering on a set S if < is a set of ordered pairs of elements in S such that:

1. For every element a in S, (a,a) is not in < (nonreflexive property).

2. For every pair of elements a, b in the set S, if (a,b) is in < then (b,a) is not in < (asymmetric property).

3. For any three elements a, b and c in the set S, if (a, b) is in < and if (b,c) is in < then (a, c) is in < (transitive property).

From this definition it can be seen that, apart from the fact that the relation < involves giving a response to pairs of objects (in some arbitrary set), it is primarily properties 1, 2, and 3 that define < . Further, it is clear that these three properties stipulate that an understanding of the relation < implies a corresponding mastery of the relation "not < " plus an ability to simultaneously consider pairs of relations.

Mathematically, the three basic types of logical-mathematical structures were defined by the Bourbaki (1948) in terms of their form of inverse and their manner of combination. Therefore, Piaget has held that psychologically operations (or relations) are not understood in isolation, but only as they relate to whole operational structures (Beth & Piaget, 1966). That is, an operation (i.e., operation, relation, or transformation) is not first learned and later assigned its properties (i.e., commutative property, associative property, etc., or reflexive property, transitive property, symmetric property, etc.). Rather, the meaning of an operation is derived from the structure of which it forms a part. For the seriation task described above, it was not until the comparison "longer than" came to be coordinated with the comparison "shorter than" that the comparison attained the status of a strict partial ordering relation. This point will be reconsidered in the next section; however, for now, the following observation should be made.

It is quite possible, that, for a specific finite set of objects, a pseudorelation between objects can be learned as simple S-R associations to pairs of objects without any accompanying understanding of the relation per se. As a trivial example of this phenomenon, one could take a set of ten ordered Cuisenaire rods (||||||||||) and teach a young child to say, "They are not equal in length" for any pair of rods which could be presented from the ordered set of ten rods. Such learning alone would not indicate an understanding of the relation ≠ (in length).[1] Rather, the child may have learned only a property of a particular set of objects. No understanding would be required of the relation per se. For the purposes of this paper, less concern will be given to a child's apprehension of S-R associations to pairs of objects in a specific set than to a child's apprehension of certain relational <u>structures</u> in a wide range of situations.

---

[1]If two identical Cuisenaire rods (A and B) are glued to a piece of paper with arrows drawn on the paper as follows (⟸⟹) and if a third identical rod, C, is used by a child to compare with rods A and B, he may say that A = C (in length) and B = C (in length), but still maintain that A ≠ B (in length).

The Formalization of Operational Structures

The preceding section posed an apparent instructional dilemma because of a phenomenon which can be called *structural integration.* Structural integration occurs when lower order concepts are brought together into a whole (i.e., structure) and when the properties of the lower concepts depend partly or entirely on the characteristics of the whole.

Structural integration occurs fairly often in mathematics. Two examples have already been given. The mastery of the set union operation +, and the ordering relation < , both involved structural integration. In order to come to an understanding of either of these individual operations, a child must consider the operation to be part of a system of operations. But this fact causes a "chicken-egg" sort of instructional dilemma. It appears that in order to master an operation, a child must master a system of operations. But in order to master a system of operations, the individual operations must be mastered.

Mathematicians can formalize a mathematical structure (e.g., define a strict partial ordering relation) by starting with certain axioms, undefined terms, or accepted rules of logic, and construct theorems and definitions on the basis of these. That is, axiomatics terminates endless regressions by beginning with undefined terms and it avoids circularity by arbitrarily choosing a starting point which has not been demonstrated.[2] Psychologically, however, one is not afforded the luxury of beginning with indefinables, axioms, or accepted rules of logic.

For example, in the case of the ordering relation < , the nonreflexive, asymmetric, and transitive properties cannot be used as self-evident concepts. Before the relation < has been coordinated with its inverse, each of these properties is repeatedly and often emphatically denied by children (Inhelder & Piaget, 1964). Even such mathematically primitive concepts as Hilbert's order axiom (If B is between A and C, then it is also between C and A) are not a priori intuitions for children until the betweenness relation has been subsumed within a system of relations (Piaget & Inhelder, 1971, p. 144).

The Genetic Construction of Operational Structures

In order to teach children a concept of redness, one can present examples and counterexamples of red objects. In a certain respect, the abstraction of operations can be achieved through a similar process. That is, operations are abstracted from many different situations in which the

[2]Gödel (1934) demonstrated the impossibility of establishing the noncontradiction of any deductive theory solely by methods borrowed from this theory or from weaker systems. The verification of the completeness and noncontradiction of a system and the independence of its axioms must be tested by the use of mathematical models. However, as soon as lower systems are subordinated to higher, only systematic wholes are guaranteed an autonomous existence (Beth & Piaget, 1966, p. 272). Referring to exactly this point, Bertrand Russell is known to have quipped, "Mathematics is the subject where we never know what we are talking about, nor if what we are saying is true."

The Formalization of Operational Structures

The preceding section posed an apparent instructional dilemma because of a phenomenon which can be called *structural integration*. Structural integration occurs when lower order concepts are brought together into a whole (i.e., structure) and when the properties of the lower concepts depend partly or entirely on the characteristics of the whole.

Structural integration occurs fairly often in mathematics. Two examples have already been given. The mastery of the set union operation +, and the ordering relation <, both involved structural integration. In order to come to an understanding of either of these individual operations, a child must consider the operation to be part of a system of operations. But this fact causes a "chicken-egg" sort of instructional dilemma. It appears that in order to master an operation, a child must master a system of operations. But in order to master a system of operations, the individual operations must be mastered.

Mathematicians can formalize a mathematical structure (e.g., define a strict partial ordering relation) by starting with certain axioms, undefined terms, or accepted rules of logic, and construct theorems and definitions on the basis of these. That is, axiomatics terminates endless regressions by beginning with undefined terms and it avoids circularity by arbitrarily choosing a starting point which has not been demonstrated.[2] Psychologically, however, one is not afforded the luxury of beginning with indefinables, axioms, or accepted rules of logic.

For example, in the case of the ordering relation <, the nonreflexive, asymmetric, and transitive properties cannot be used as self-evident concepts. Before the relation < has been coordinated with its inverse, each of these properties is repeatedly and often emphatically denied by children (Inhelder & Piaget, 1964). Even such mathematically primitive concepts as Hilbert's order axiom (If B is between A and C, then it is also between C and A) are not a priori intuitions for children until the betweenness relation has been subsumed within a system of relations (Piaget & Inhelder, 1971, p. 144).

The Genetic Construction of Operational Structures

In order to teach children a concept of redness, one can present examples and counterexamples of red objects. In a certain respect, the abstraction of operations can be achieved through a similar process. That is, operations are abstracted from many different situations in which the

[2]Gödel (1934) demonstrated the impossibility of establishing the noncontradiction of any deductive theory solely by methods borrowed from this theory or from weaker systems. The verification of the completeness and noncontradiction of a system and the independence of its axioms must be tested by the use of mathematical models. However, as soon as lower systems are subordinated to higher, only systematic wholes are guaranteed an autonomous existence (Beth & Piaget, 1966, p. 272). Referring to exactly this point, Bertrand Russell is known to have quipped, "Mathematics is the subject where we never know what we are talking about, nor if what we are saying is true."

operation occurs. Piaget has held that these situations in which operations occur are the child's actions. Logical-mathematical knowledge is seen as beginning, not with an awareness of self (a priori intuitions) or of things, but with the coordination and recognition of their interactions. That is, an operation is an internalized *scheme* of interactions, where the scheme of a set of actions is their common operational essence.

> *The scheme of an action is, by definition, the structured group of the generalizable characteristics of this action, that is, those which allow the repetition of the same action or its application to a new content. . . . This is why such schemes have a completely general significance and are not characteristic merely of one or another of the actions of a single individual.* (Beth & Piaget, 1966, p. 235)

For example, in order to teach a child the relation < , he can be given a variety of experiences ordering many kinds of objects according to various criteria. To expect a child to abstract the relation < by working with only one set of objects (e.g., Cuisenaire rods) would be as unlikely as expecting the child to abstract redness by showing him only one red object.

Thus, internalizing schemes of actions means abstracting the common operational essence from a number of isomorphic interactions. In the case of the seriation operation, one might give the child the following types of experiences:

1. put Cuisenaire rods in a row according to length,
2. put dowel rods in order according to length,
3. put cylinders in a row according to height,
4. put cylinders in a row according to diameter,
5. put circles in a pile according to diameter,
6. put cubes in a row according to size,
7. make a tower of cubes according to size,
8. put spheres in order according to size,
9. put spheres in order according to color,
10. put sandpaper in a pile according to roughness.

None of the materials above inherently embody the relation < . Rather, each set of materials can be used to coordinate the relevant scheme of interactions. These various sets of materials can then be used to help the child abstract the relevant system of operations. However, the abstraction which takes place is not from the objects per se, but from the systems of interactions that were coordinated using the objects.

Reflexive Abstraction

In the previous section certain similarities which exist between teaching children a concept such as redness through experiences with objects and teaching children an operation through experiences performing actions on objects were pointed out. However, there are also certain dissimilarities between these two types of abstractions. In order to abstract a concept of redness from a set of objects, the child simply needs to isolate the relevant property. However, as long as the single interaction is isolated, it can have little significance to the child as the archetype of an operation. To abstract operations from one's own actions consists not simply of taking note of individual isolated interactions, it requires the reconstruction of these actions on a higher plane. Individual interactions gradually take on new significance (reflexive abstraction) as they are modified by being treated as part of a whole operational structure.

Reversibility

The key to the emergence of a whole system of schemes of interactions is the appearance of the inverse to the given scheme. This reversibility phenomenon is attained when the child exhibits a recognition of the fact that the combined application of a scheme followed by its inverse is equivalent to the identity scheme. According to this definition reversibility implies not only the emergence of the inverse scheme but also the identity scheme and combination of pairs of these.[3] Thus, the attainment of reversibility implies the existence of an operational system which in turn elevates the scheme to the status of an operation as part of the structured whole.

For Piaget, an operation is an internalized scheme of interactions that is reversible and that is dependent on other schemes with which it forms an operational system characterized by laws that apply to the whole structure (Beth & Piaget, 1966, p. 234). Further, he has held that the simplest such structures include not just the original scheme of actions, but also at least its inverse, identity schemes, and combination of pairs of these.

Genetic Circularity

It should be clear from the above description that the evolution of operational structures would not be conceived as beginning with individual isolated operations which are successively linked together. Rather, the evolution of structures of operations would be visualized as occurring simultaneously with the evolution of the operations that the structure subsumes. Thus, both the structure and its operations simultaneously crystalize out of a system of schemes of actions as it becomes progressively coordinated (genetic circularity).

While the coordination of a system of schemes of actions is achieved progressively, its completion is marked by a momentary acceleration in this construction as the child shifts to a qualitatively higher level of thought. As a result of this reorganization, new self-evidence typically appears

---

[3]One could easily argue that the emergence of the identity scheme is equally as important as the emergence of the inverse and choose to call this event the child's recognition of Identity (Berlyne, 1965). Or one could argue that the child's ability to combine pairs of schemes is the significant event (Lunzer, 1960a, p. 32) and call the event Combination. The choice of terms seems somewhat arbitrary, however, in the sense that Reversibility, Identity, and Combination should each be implied no matter which of the three terms one wishes to emphasize.

with regard to concepts whose definitions depend upon the application of the given structure. Thus, certain operational concepts (such as the concept of a series, or class) and certain properties (such as transitivity) arise out of structured wholes of operations, the completion of which explains the necessity of its elements insofar as their meanings are dependent on that whole.

Piaget's Groupings

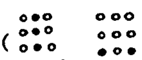
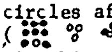
Because of the fact that a formalization of the most elementary operational structures which are mastered by children are globally similar to a mathematical group, Piaget has coined the word "grouping" to refer to such structures. The concept of a grouping is believed to be useful (1) due to the ability of children to internalize any particular action that is included within a given scheme of action at "about" the same time and, (2) since the internalization of a scheme of action automatically implies the internalization of a structure of internalized actions, the simplest of which includes the scheme, its inverse, identity schemes, and combinations of pairs of these.

It should be emphasized that a grouping is not some sort of "a priori" cognitive structure (à la Gestalt psychology) which imposes itself on the thought of a child. There is no more reason to attribute a priori existence to Piaget's groupings in the minds of children than there is to attribute a priori existence to the Pythagorean Theorem in the mind of Pythagoras. The theorem was not a form into which Pythagoras' experiences fell. It was a necessary consequence of the progressive organization of Pythagoras' experiences. There is an important psychological distinction between a form that is assumed to exist a priori in a child's mind and a form that is a product of equilibration (i.e., progressive organization) and which could not have developed otherwise.

Cognitive Characteristics of Preoperational Children

Children who have not yet mastered a given system of operations are characterized by at least the following cognitive characteristics. syncretic thinking, centering, and fixed state thinking.

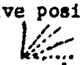
Syncretic Thinking

Children who are able to copy a 3 x 3 array of circles (), and who have stated that each of the two arrays contains nine circles, are often convinced that the two arrays no longer contain the same number of circles after one of the two arrays has been partitioned into three clusters (). Children have difficulties coordinating part-whole relationships within numerical aggregates. When attention is drawn to a numerical whole, the parts (or units) are cognitively neglected. When attention is directed toward component parts, the whole is often cognitively destroyed.

That is, children tend to view sets of objects *syncretically*, or as a global unanalyzed whole, rather than analytically.

Gast (1957) has shown that the ability to determine the cardinality of a set depends on the homogeneity of its items. He found an initial stage in which virtually complete homogeneity of the elements is required; a second in which perceptual diversity is possible within certain limits of qualitative resemblance; and a final state in which the objects may belong to several disjunctive classes (also see Piaget, 1952; Dodwell, 1962; and Elkind, 1964). The concept of a unit is by no means an a priori intuition for young children.

Fixed State Thinking

Preoperational children tend to focus on successive states of an object or set of objects rather than on the operations that connect these states. For example, if a kindergarten child is asked to represent (by drawings, by gestures, by multiple choice selections from pictures) the successive positions occupied by a stick in falling from vertical to horizontal (), the task proves to be surprisingly difficult.

While kindergarteners are usually able to represent the beginning vertical position and the final horizontal position, intermediate positions often present great difficulties. Young children may not only fail to represent intermediate positions correctly, they may even fail to recognize a correct representation when it is shown to them (Piaget & Inhelder, 1971). Thus, preoperational children seem unable to integrate a series of states of an object into a continuous whole—or a transformation.

As another example consider two identical transparent bowls one of which contains six beads and the other five and two identical opaque boxes in front of each bowl (Figure 1). If the beads are taken from the bowls and placed one for one into the boxes until each box contains five beads, young children may believe that there are not the same number of beads in

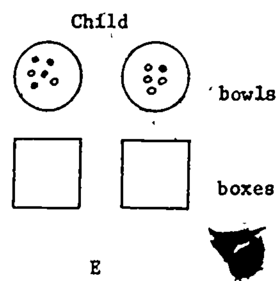


Figure 1

the two boxes, indicating that a yellow bead is left in one bowl while all of the blue beads have been used. Further, if the child is told that there are five beads in one of the boxes and is asked to guess how many beads are in the other, he will typically answer in accordance with his judgment of equality or inequality. If he had previously stated that the two sets were unequal, he will likely guess almost any number other than five. Thus, young children seem to believe that the cardinality of a set may somehow be affected by the source of its elements.

Preoperational children tend to base their judgment solely on isolated configurations that are before them at any given moment. This cognitive preference was reflected in the above task in that children typically make their incorrect numerical judgment based on the only number-relevant information that was available in the final state of the task situation--the source of the objects.

Centering

Another task can be posed which illustrates one of the cognitive characteristics that accompanies thinking in terms of unconnected fixed states. Place two cylindrical glasses in front of identical boxes, each of which contains thirty half-inch beads (Figure 2). The child is directed to take a blue bead in one hand, a yellow bead in the other and to put the beads into the two cylindrical glasses at exactly the same time. After ten beads have been put into each of the two glasses, the child may believe

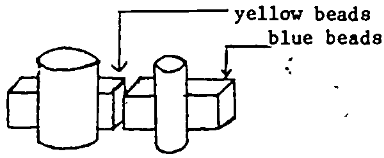


Figure 2

that there are more beads in the glass where the beads reach a higher level.

Children not only focus their attention on momentary conditions of an object or set of objects, they also *center* on only the most salient perceptual features of a given configuration. That is, they may notice height but neglect thickness.

The Concept of Number

It has been argued that it is typical for logical-mathematical concepts to be operational concepts. Nonetheless, the involvement of operations in the formation of many mathematical concepts is much less obvious than in the two examples cited thus far. In particular, the operational nature of early number concepts has been contested throughout the history of mathematics education. One of the most eloquent arguments against nonoperational points of view concerning the origin of number concepts has been given by Dewey (McLellan and Dewey, 1914), whose views are summarized in the following statements:

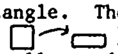
> Number *is not (psychologically) got* from *things, it is put* into *them (p. 61).*
>
> . . . *This abstraction is* complex, *involving two factors: the difference that makes the individuality of each object must be noted, and yet the different individuals must be grasped as one --a sum (p. 25).*

Thus, in considering a group of seven red circles, Dewey would contend that the concept "redness" is a qualitatively different sort of concept than the concept of "sevenness." Whereas redness can be directly perceived in the circles, sevenness only becomes a property of the circles due to operations that are performed on the circles.

Modern mathematics educators have often cited children's responses to Piagetian number conservation tasks in order to help verify the point of view expressed by Dewey. Nonetheless, universal acceptance has not been given to any one of the various possible interpretations which exists concerning the relationship between a child's understanding of number concepts and his ability to respond correctly to conservations of number tasks. In the following two sections, conservation-like tasks will be given which will help to establish an interpretation of the significance of conservation tasks. In addition, the examples may help to clarify the way children come to master elementary number concepts.

Qualitative Cognitive Growth

Implicitly taking the position that correct conservation responses are largely unrelated to the child's level of understanding of the concept of number, certain psychologists and educators have attempted to explain the mastery of conservation tasks by children in terms of their gradual mastery of the fact that a spatially displaced set of objects can be returned to their original positions (or slight modifications of this argument). It seems likely that an understanding of empirical return is a necessary (but not sufficient) condition for mastery of conservation tasks. However, the insufficiency of this explanation is illustrated by the following example.

Suppose a circular string of circumference four inches is put in the shape of a square (). The child is told that the string is a fence and that a cow can eat the grass inside the fence. Then the string is changed into the shape of a 1-1/2 inch by 1/2 inch rectangle. The child is asked if the cow still has the same amount to eat ().

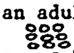
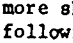
Understanding of empirical return is an almost equally good explanation of conservation of area in the above task, in spite of the fact that area is not conserved for this task as can be seen if the string continues to be transformed into two line segments of length two inches ().

Kindergarten children often respond correctly that the area diminishes as the square becomes more rectangular. However, to assume that these children understand the concept of area would be incorrect. Further questioning may indicate that those children who responded correctly were basing their judgments on the height of the rectangles. (Incorrect answers usually focus on the width. Only seldom will a child judge the areas to be the same.) Thus, children often respond correctly by basing their judgments on the wrong information.

Children older than five years of age (and in fact many adults) often assert that the area remains the same. Further, they may maintain this conviction almost until the area disappears in the limiting case. Such adults certainly do not understand less than the average kindergarten child. They have shifted to a *qualitatively* higher level of thought which brings with it new factors as sources for incorrect judgment. A fundamental fact which Piaget's research makes abundantly clear is that cognitive growth from birth to maturity does not simply get qualitatively better and better; qualitative reorganizations also occur. His theory addresses itself to these qualitative changes. Learning is not simply a matter of associating right answers to questions.

If one's emphasis is on a child's understanding of the concept of number, it is crucial to be able to account for the differences between Piaget's conservation of number tasks and the area task described above. That is, it is important to be able to explain how children come to understand that *number* is invariant under simple displacements. Too often invariance (or conservation) has been treated as a unitary sort of concept, as though invariance of what (i.e., number, mass, volume, area, etc.) were relatively unimportant.

Rejectable Explanations of the Initial Attainment of an Operational Concept of Number

With respect to logical-mathematical knowledge, Piaget has considered the most relevant aspects of intelligence to be, not what a child perceives, but the rules of organization which the child gradually develops in order to control and use the information he receives. Although an adult may feel that he perceives "nineness" in a 3 x 3 array of circles (), he may be more skeptical regarding the purely perceptual origin of "nineness" in the following configuration (). The sensation of perceiving "nineness" appears to be similar to what happens when one looks at a hidden picture puzzle. Once the picture is distinguished, it is difficult to realize how it had ever been disguised (Bruner, 1968, Ch. 5).[4]

Many examples could be given which would bear witness to the fact that any concept of number which does not involve at least the operations of giving an order to objects which previously had no order (seriation), identifying in some sense objects which are not in fact identical (classification), and coordinating part-whole relationships in order to grasp the concept of units can involve only the most rudimentary and superficial concept of number. Until children have mastered these elementary operations, they not only fail to realize that number is invariant under simple displacements, they also deny each of the properties that define the concept of number. For instance, tasks can be posed in which nonconservers will deny the validity of Peano's Axioms of (1) that a numerical whole is equal to the sum of its parts, (2) that addition is commutative (or associative), (3) the existence of an identify element (or inverses), and (4) the relationship between cardinal and ordinal numbers (Piaget, 1952). Such tasks can be used to determine progressively reduced developmental levels at which a nonoperational concept of number could exist.

Having reduced the level of development at which a nonoperational concept of number could exist, the question remains whether an early operational number concept must ultimately, or at some still lower level, evolve out of some sort of nonoperational concept. One might still hold that some nonoperational concept (although exceedingly rudimentary) is actually

[4]Wohlwill (1968) has summarized some of the distinctions between perception and conception as seen by the Gestalt school (Kohler, Wertheimer, Bruner, Brunswick, and Piaget).

a first approximation to a later, more refined concept. On the other hand, it may equally well be that any nonoperational concept of "number," far from being a first approximation to a more mathematically viable concept of number, is actually a detriment to later learning. To clarify this possibility, consider the following analogy.

A child who could select a red crayon from his box of crayons would likely be considered to have attained a primitive concept of redness. To be sure, the concept would be quite elementary to a spectrometrist, who must analyze incoming light from far away stars. Nonetheless, under limited conditions, the concept would likely suffice as a first approximation if it did not deny any of the properties that characterize a later, more sophisticated notion of redness.

Conversely, a concept which could not in some sense be construed as being a first approximation to a later, more sophisticated concept would not be considered to be a concept of redness. Thus, if the child's red crayons were broken in half and he were taught to recognize it only by its shorter length, the learning which would accrue would not be a concept of redness. Even though the child might always be able to select the red crayon from his box upon command, he might be responding to length cues rather than to color cues. Such a training procedure would be foolish, of course, since redness is at least as easy to teach a child as shortness. It does little good to trick a child into giving a correct response to erroneous cues and it does little good to trick him into correct responses which he does not understand. However, while these maxims seem to be so much a matter of common sense concerning the concept of redness, they are commonly violated concerning children's instruction pertaining to number concepts. Children are often tricked into giving correct arithmetic responses which they do not understand and/or into giving correct responses to erroneous cues.

For example, while a child can often be tricked into giving the correct numerical responses to the following arrays of objects (· : ∴ :: ⁙, etc.), the concept that the child may be learning may not be a concept of number. The essence of making numerical judgments involves learning to avoid making judgments on the basis of shape or pattern. A training procedure based on standard dot patterns may encourage preoperational children toward tendency to judge numerical wholes on the basis of gross configuration (or area covered) rather than on appropriate cues.

Dot patterns, Cuisenaire rods, counting discs, and arithmetic blocks can each be useful models in order to help children come to an understanding of number concepts. However, even if a dot pattern (⁙ ⁙) can eventually be used as a model to represent the number 10, it is important to remember that it is a constructed representation. That is, the model only comes to embody the number 10 after certain systems of operations have been coordinated relative to the model. Until a child has coordinated these operations into elementary systems, his thinking will tend to be fixed and syncretic with respect to tasks that are characterized by the structure, and he will tend to center on only one aspect of models that are presented.

Initially Purely Verbal or Symbolic Concepts of Number

The preoperational thinking of young children seems to be so different from that of an adult that for the child many adult-like words and responses must mean something very different from the meaning an adult assigns to them. It is difficult to say exactly what the statement 5 + 3 = 8 may mean to a little child whose thinking is characterized by fixed thinking, syncretic thinking, and centering; but it is well known that even without proper understanding, children are quite capable of memorizing large quantities of verbal or symbolic material. Thus, misunderstandings are often not detected until an entire facade of ill-conceived notions collapses. In early elementary school arithmetic, this collapse typically occurs when children reach regrouping concepts involving "borrowing" and "carrying."[5]

The improvement of langauge may aid in the acquisition of an operational concept, if the activation of language can facilitate the coordination of operations and enable the child to be less dominated by perceptual forces (Bruner, 1964). Nonetheless, appeals to an initially purely verbal-symbolic concept to explain the development of an operational number concept is insufficient. Piaget (Ripple & Rockcastle, 1964) has stated:

> *Words are probably no short-cut to a better understanding . . . . The level of understanding seems to modify the language that is used, rather than vice versa. . . . Mainly language serves to translate what is already understood; or else language may even present a danger if it is used to introduce an idea which is not yet accessible (p. 5).*

The fundamental problem appears to be to determine to what underlying concepts the language and symbols that children use are being attached (Bruner, Olver & Greenfield, 1966, p. 47).

The Genetic Development of the Concept of Number

On the basis of parsimony, Piaget's description of the development of number concepts is quite pleasing since an operational number concept need not be assumed to evolve out of some sort of nonoperational concept. Piagetians (e.g., Piaget, 1952; Inhelder & Piaget, 1964) have furnished an impressive quantity of data to substantiate the hypothesis that elementary number concepts (the assignment of numerals to sets whose elements are regarded as classed and ordered) develop in parallel to, and as a synthesis of, the development of elementary classification and seriation concepts. This appears to be so since the intellectual coordinations involved in forming series and classes are also involved in forming seriated classes (i.e., numbers). However, Piaget's analysis of the cognitive evolution of number and other logical-mathematical concepts relies heavily on the psychological viability of analyzing, ordering, and equating concepts (or tasks) on the basis of their underlying operational structures.

---

[5]Roughhead & Scandura (1968) and Brownell & Chazal (1935) have reported findings which indicate that rote verbal or symbolic learning may actually cue out the kind of reorganizing activity which seems necessary for a child to come to an operational understanding of number.

Ordering and Equating Operational Structures

For Piaget, cognitive growth is viewed as a process of gradually coordinating systems of operations through the dual process of assimilation (i.e., inner organization) and accommodation(i.e., outer adaptation).

As the child's perceptual activities become coordinated, he becomes cognizant of more features of what is perceptually before him at any given moment (i.e., less centering, syncretic thinking, and fixed state thinking). This increased awareness demands greater coordination, which in turn produces still more analysis of perceptual givens. In short, what we have is an activity that organizes reality while coordinating its own functioning. The tendency in adaptation is constantly in the direction of greater equilibrium of the functioning structure in the face of external disturbances and demands for internal consistency (i.e., coordination) (Piaget, 1960). Further, each relative equilibrium state carries with it the ability to detect new sources of disequilibrium and hence the seeds of its own destruction. What we have here is a sort of concrete analogue to the fact established by Gödel (1934) concerning the impossibility of establishing the noncontradiction of a deductive system solely by methods borrowed from this system or weaker systems.

This progress toward greater equilibrium, together with the proposition that operations exist psychologically only within structured operational wholes, yields a basis for ordering and equating operational concepts (or tasks). As a trivial illustration, one would expect that the task of putting cylinders (which vary according to height and diameter) into 4 x 4

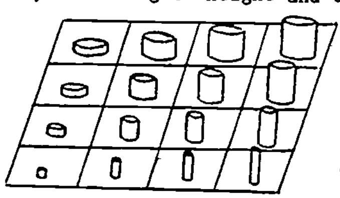


Figure 3

matrix would be mastered no sooner than the task of putting ten dowel rods in order according to length. This is because the operational structure that characterizes the dowel rods task is included within the structure that describes the task involving a matrix of cylinders (Figure 3).

Intra-individual Variability

While Piagetians have amassed a large amount of data to support the contention that tasks with isomorphic operational structures are mastered at "about" the same time, it is also well known that intra-individual variability commonly occurs concerning a child's ability to perform tasks which are characterized by a single operational structure (Piaget, 1971, p. 173).

As an example of the phenomenon of intra-individual variability, consider the following tasks. Six pennies are placed in front of a kindergarten child, and ten pennies are placed in front of an adult. The child is then asked to "Make it so we both have the same number of pennies." Another similar task can be posed using small one-inch cubes instead of pennies. The problem is markedly more difficult using pennies than using one-inch cubes. Using cubes, the problem is commonly solved by kindergarteners following the sequence of steps illustrated in Figure 4. The child is aided in making the two groups equal in number by being able to make them the same shape. Using pennies, the child is forced to be more analytic in his consideration of the two groups. Therefore, if one were to consider a set of tasks all of which are characterized by the same

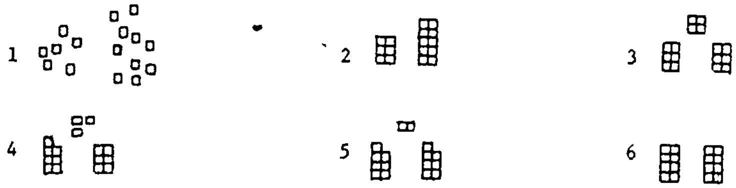


Figure 4

system of operations, the tasks would vary somewhat in difficulty due to the relative involvement of factors such as: syncretic thinking, fixed state thinking, and centering.

Ordering and Equating Tasks

The intra-individual variability, which is part of Piaget's theory concerning a child's ability to perform concrete tasks which are characterized by a single operational structure, has caused understandable confusion among those who would interpret this theory. This confusion seems to have developed, at least in part, because of a common failure to distinguish between the invariant sequential mastery of various tasks which involve the application of these operations. To illustrate this distinction, consider the following situation.

Suppose that tasks T_1 and T_1' both involve the application of an operational structure S_1. The theory predicts that tasks T_1 and T_1' (and all other tasks characterized by operational structures S_1) should be mastered at "about" the same time, subject to a certain amount of intra-individual variability, and subject to certain side conditions of equivalence between the two tasks.[6] Hence, one might visualize a given child's mastery of all tasks that are characterized by a particular operational structure, as in Figure 5.

In particular, then, it would only in general be true that tasks T_1 and T_1' are mastered at the same time. For instance, it could happen that task T_1' was mastered before T_1.

Now suppose that another task T_2 was found to be characterized by an operational structure S_2, and further suppose that structure S_2 includes (or subsumes) S_1. On the basis of the subsumption of structure S_1 by S_2,

[6]That is, the two tasks would have to be equally facilitating. They would have to be relatively equivalent concerning the degree to which they require the child to decenter, be more flexible, be more analytic (i.e., less syncretic), etc.

one could conclude that structure S_2 would be mastered by a child no sooner than structure S_1. Nonetheless, this fact would not necessarily imply that

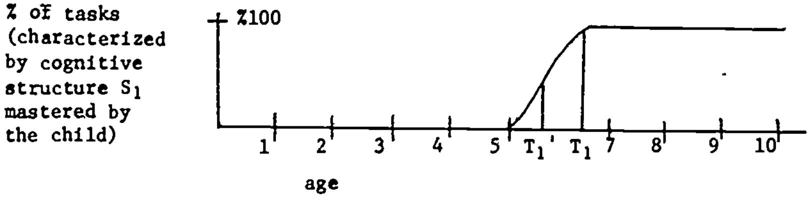


Figure 5

task T_2 would be mastered no sooner than task T_1. In fact, the situation could occur as illustrated in Figure 6. That is, although it is in general true that tasks characterized by structure S_2 are mastered no sooner than tasks characterized by structure S_1, this might not be the case with respect to particular tasks T_2 and T_1. It could happen that task T_2 would be mastered

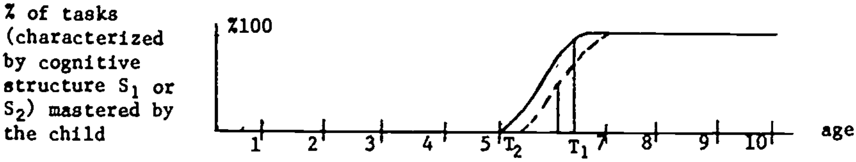


Figure 6

before task T_1.

Lunzer's Hypothesis

Lunzer (1960a, pp. 30-32; 1960b) has attempted to reconcile the fact of intra-individual variability with the quantity of evidence that Piagetians have produced showing that children attain the general ability to internalize *all* logical-mathematical action schemes into reversible structures over a relatively short period of time. That is, children usually master all three types of logical-mathematical groupings at about six or seven years of age--subject to variations due to factors such as experience, social transmission, and equilibration (Piaget, 1964a).

Lunzer has suggested that the crucial step that is taken by children at about six or seven years of age is when they become capable of making two judgments simultaneously, and this ushers in the beginning of Piaget's period of Concrete Operations.

It is, no doubt, quite true that a generalized ability to make two simultaneous judgments is a prime factor accounting for the great cognitive reorganization which takes place in children at about six or seven years of age. In fact, this seems to be only another way of defining Piaget's concept of reversibility. What is for Piaget the coordination of an operation with its inverse is for Lunzer the ability to make two simultaneous judgments. However, Piaget has believed that it is fruitful to distinguish at least three types of logical-mathematical operations (Beth & Piaget, 1966, p. 186). In terms of simultaneous judgments, these operations would perhaps be distinguished as follows.

1. Operations involve coordinating two properties of a set of objects.
2. Relations involve coordinating two comparisons between objects.
3. Transformations involve coordinating two perceptions of an object.

The main factor that enables all three types of groupings to be mastered during the same general age range may well be the fact that each involves the coordination of an operation with its inverse. Nonetheless, the three basic types of logical-mathematical operations may retain certain unique characteristics that enable them to be distinguished from one another as distinct psychological entities (Elkind, 1964).

It is only in such special instances as the synthesis of the classification and seriation groupings to form the number group that Piaget has hypothesized functional interdependence between concrete operational structures. The relationship between the number group and the classification or seriation groupings is predicted to be closer than the relationship between the classification and seriation groupings. The relationships between specific actions within a given scheme are closer than the relationships between groupings that are formed from distinct schemes.

The Pilot Study

A theory which hypothesizes the invariant sequential mastery of certain operational structures while allowing for intra-individual variability concerning a child's ability to perform tasks which are characterized by these structures has remained a source of controversy. For example, certain psychologists (e.g., Kohnstamm, 1967) have asserted that the fact of intra-individual variability renders meaningless the practice of ordering and equating tasks on the basis of underlying operational structure. Such criticisms suggested a pilot study, the primary experimental objective of which was to investigate the interdependent development of classification. seriation, and number concepts.

While helping to confirm the psychological viability of analyzing, ordering, and equating tasks on the basis of their operational structures, the pilot study was to serve the dual function of furnishing the theoretical scaffolding which would be necessary for structuring and interpreting a transfer of training study. Toward these ends, three parallel sequences of Piagetian tasks were obtained (denoted by S1, S2, . . ., S7; N1, N2 . . ., N6; and C0, C1, C2,. . ., C6) which were graded in difficulty and which pertained to seriation, number, and classification respectively. Each of these three sequences was determined by ordering tasks according to the theory outlined in preceding sections and selecting those tasks which would, in fact, exhibit a relatively invariant sequential mastery. For example, the seriation tasks were related in such a way that the probability would be small (<15%) that a child would be able to correctly respond to task

S(n + 1) before he is able to respond to task S(n); and similarly for the other sequences of tasks. Once formulated, these three sequences were to be used to investigate the interdependent development between concept areas.

The easiest task in the seriation series involved copying a circular string of beads (S1) or copying a string of beads in the inverse direction (S2). Progressively more difficult seriation tasks involved reconstructing a set of ten ordered Cuisenaire rods (S3), putting ten dowel rods in order according to length by trial and error (S4), or without trial and error (S5), inserting dowel rods into a completed series (S6), and reconstructing a 4 x 4 matrix of cylinders (S7). The simplest classification task involved producing "nongraphic collections" (see Inhelder and Piaget, 1964, p. 19-20, for an explanation of classification by graphic collections) when attempting to classify objects within a set of yellow cylinders, yellow cones, green cubes, and green pyramids (C1). More difficult classification tasks involved anticipating criteria for exhaustively subdividing sets of objects (C2), repartitioning sets of objects according to differing criteria (C3), reconstructing a 5 x 5 classification matrix (C4), and hierarchically classifying a set of objects (C5). The most difficult classification task (C6) was a quantitative inclusion task (Inhelder & Piaget, 1969, Chapter IV). The simplest number tasks involved copying a row of seven circles (N1) and partitioning a set of sixteen pennies into four equal sets (N2). More difficult number tasks involved equalizing a set of ten pennies and six pennies by taking pennies from the larger set (N3) and copying a 3 x 4 array of circles (N4). The most difficult number tasks were four distinct types of number conservation tasks.

After a child had copied a row of seven red circles (●●●●●●● / ○○○○○○○), conservation task N5a required the child to realize that his row still had the same number of circles as the model row after the circles in the model row had been pushed closer together (○ ●○●○●○●○ ○ ○). Conservation task N5b tested the child's understanding of the fact that the number of circles in a 3 x 4 array (○○○○ / ○○○○ / ○○○○) does not change when the circles are regrouped into three 2 x 2 arrays. Task N6 involved two subtasks, one in which beads were placed into two separate and different shaped glasses by the child (N6a), and the other in which six beads were placed into two separate identical red cardboard boxes in a one-to-one fashion where the beads came from piles of beads of different cardinality (N6b).

Procedure

Each of the three sequences of tasks was administered individually to each of 160 kindergarten children during the last month of the 1969-1970 school year. The children participating in these studies represented rather typical small town Indiana communities. Although I.Q. test scores were not available for kindergarten children in the schools involved, typical achievement test and I.Q. test performance of older children in the participating schools were about average for the state of Indiana. Severely mentally handicapped children had been identified (and placed in special classes which were not used in the study) through the combined use of a Bender Perceptual Development Test, a Boehm Test of Basic Concepts, and a Draw-a-Man General Intelligence Classification Test. Scores for the children used in the pilot study sample on each of these tests were distributed from below average to superior for their age group. The ages of the children (83 boys, 77 girls) ranged from 5 years 4 months to 6 years 7 months. Each of the 160 children took the task batteries on three successive days. Each of the three batteries of tasks required from 10 to 20 minutes to complete. The order in which the test batteries were administered was varied. One-sixth of the children were chosen at random to take the seriation battery of tasks on the first day, the number battery on the second day, and the classification battery on the third day. Similarly, one-sixth of the children took the task batteries in one of the other six permutations of the order: seriation, number, and classification.

Results

Tasks on the pilot study batteries were each scored on a pass-fail basis. A summary of the results of the pilot study are recorded in Figure 8, which should be read as illustrated in Figure 7. Figure 7 indicates that of the 160 children who participated in the pilot study, 45 responded correctly to task T_1 but not to task T_2, and 5 responded correctly to task T_2 but not to task T_1. Other children either missed both tasks or responded correctly to both tasks. Thus, Figure 7 reveals that, on the basis of a sample of 50 children who responded incorrectly to exactly one task, the chances are approximately 90 percent, i.e., $90\% = 45/(45 + 5) = 45/50$ that a given child will correctly respond to task T_1 before task T_2.

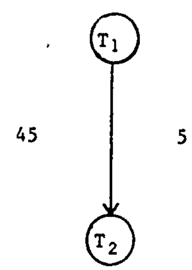
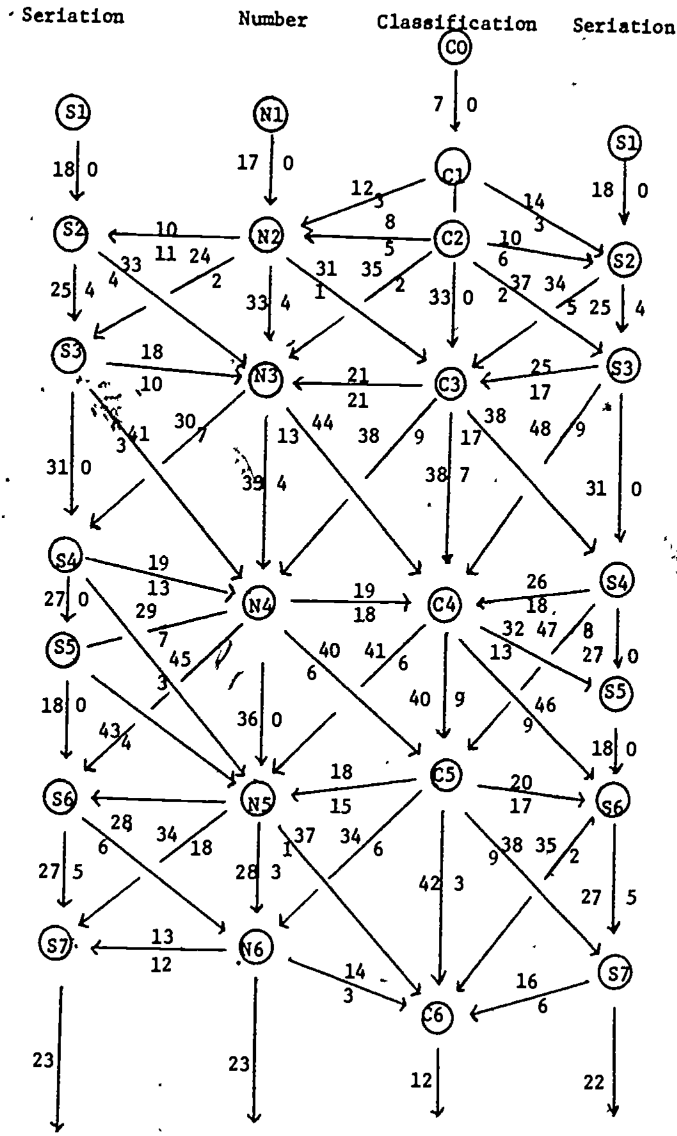


Figure 7

The tasks are related in difficulty approximately as illustrated in Figure 8. The easiest tasks are at the top of the figure and children appear to proceed in a parallel fashion through each of these three series of tasks. That is, tasks which are at approximately the same level in Figure 8 are comparable in difficulty, whereas tasks which are at different levels differ significantly as to degree of difficulty.

The existence of a high correlation between cognitive development in these three concept areas is also evidenced by the following information. By assigning each child an ordered three-tuple (x,y,z), where x, y, and z correspond to the child's scores on the seriation, number and classification batteries respectively, scatter diagrams and Pearson's product-moment coefficients were obtained.

Figure 8


Seriation
Number
Classification
Seriation
C0
S1
N1
C1
S1
S2
N2
C2
S2
S3
N3
C3
S3
S4
N4
C4
S4
S5
S5
S6
N5
C5
S6
S7
N6
C6
S7

Although the correlation coefficients between these three series of tasks were high ($r_{ns} = .695$, $\sigma_{r_{ns}} = .041$; $r_{nc} = .650$, $\sigma_{r_{nc}} = .046$; and $r_{sc} = .609$, $\sigma_{r_{sc}} = .050$), the nature of the relationship between the three areas is by no means established. In particular, a causal connection between these series based on transfer of learning with respect to underlying operational structures (as opposed to simply a high probabilistic correlation) has not been established. However, several facts are of interest in this regard from the information that was obtained.

It was noted that correlation coefficients $r_{ns} = .695$ ($\sigma_{r_{ns}} = .041$) and $r_{nc} = .650$ ($\sigma_{r_{nc}} = .046$) were both greater than the correlation coefficient $r_{sc} = .609$ ($\sigma_{r_{sc}} = .050$). Such a result, if significant statistically, would be consistent with Piaget's hypothesis of closer operational ties between number and either of the other two concept areas than between seriation concepts and classification concepts. However, (with the possible exception of the difference between r_{ns} and r_{sc}) the above results did not attain statistical sifnificance and were not considered as evidence confirming any theoretical position. To obtain more information concerning this hypothesis, scatter diagrams were plotted recording the following information:

1. The scores on the Number Tasks versus the sum of the scores on the Seriation and Classification Tasks.
2. The scores on the Seriation Task versus the sum of the scores on the Number and Classification Tasks.
3. The scores on the Classification Tasks versus the sum of the scores on the Seriation and Number Tasks.

The respective correlation coefficients which were obtained from these diagrams were: $r_n = .809$ ($\sigma_{r_n} = .027$), $r_s = .751$ ($\sigma_{r_s} = .035$), $r_c = .724$ ($\sigma_{r_c} = .038$).

From this information it is possible to determine that, for the children and tasks used in this study, the chances are 95 percent better of predicting performance on the number tasks on the basis of the sum of the scores on the seriation and classification tasks than of predicting performance on either of the other series of tasks using a similar method. Apparently, there was a tendency for scores on the number task sequence to lie "between" the scores on the remaining two sequences.

Although the above fact is an interesting result which appears to be consistent with Piaget's position, it is considered to be more of an indication of an issue which requires further research than a piece of data which either confirms or disproves any given position. Further, the heart of the issue concerning the psychological viability of ordering and equating tasks on the basis of underlying operational structures seems to be not so much a question of how much intra-individual variability is allowable as it is a question of whether significant transfer of learning is possible between tasks which are characterized by isomorphic operational structures.

The Training Study

The key issue toward which the training study was to be directed was to determine whether transfer of learning could be induced between tasks involving similar underlying operational structures. Since Piaget has predicted number to be a synthesis of seriation and classification operations, the decision was made to try to induce number concept learning through teaching seriation and classification concepts.[7] It was hoped that the seriation and classification concepts would bear a similarity to the number concepts only due to underlying operational structure. However, Lunzer's hypothesis suggested that precautions should be taken to demonstrate that if transfer was obtained it was not due only to the ability to make two judgments simultaneously. Toward this end, a short study was conducted to determine two tasks which involve only transformations and which would be roughly equivalent in difficulty to tasks N5 and N6 from the pilot study. If such tasks could be found, it was hypothesized that transfer from learning classification and seriation operations would produce gains in the understanding of number concepts, while very little gain would be made concerning spatial concepts involving only transformations.

The two tasks which were selected involving only spatial transformations were Task T6: Piaget's three mountain problem concerning the child's ability to accurately conceive of points of view other than his own (Piaget & Inhelder, 1967, Chapter VIII), and Task T7: Piaget's task dealing with horizontal axes relative to the water level in an inclined bottle (Piaget & Inhelder, 1967, Chapter VIII). Tasks N5, N6, T6 and T7 were administered individually to 100 kindergarten children (50 boys, 50 girls). The results of the study are shown in Figure 9.

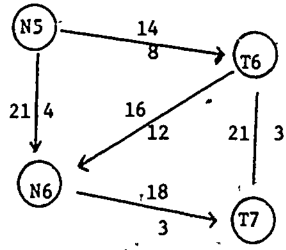


Figure 9

[7]Training studies such as those conducted by Sigel, Roeper, and Hooper (1966), Churchill (1958a), and Lasry (1969) offered hope that such training might be possible.

Subjects for the Training Study

The results of the pilot study were used to select twenty kindergarten children who were matched according to cognitive development relative to seriation, classification, and number concepts. That is, children were selected who could correctly respond to tasks S1, N2, and C1, but who responded incorrectly to task S3, N3, and C3, and who could thus be assumed to be incapable of responding correctly to more difficult tasks in any of the three sequences of tasks.[8] The children participating in the training study were all kindergarten children from one of the schools that had been used the previous year in the pilot study.

The above twenty children were divided into two groups which were equivalent in age (range: 5 years 3 months to 6 years 2 months), boy-girl distribution, performance on tasks S2 and C2, and performance on the following three tests. Bender Perceptual Development Test, Boehm Test of Basic Concepts, and the Draw-a-Man-General Intelligence Classification Test.

Procedure

The training study involved two and a half weeks between Thanksgiving and Christmas vacations. Training sessions involved groups of five children each and lasted about a half hour each school day. The training group participated in laboratory type sessions to be described in the following section. These training sessions aimed at teaching children all and only the operations which were involved in the tasks of the classification sequence or the seriation sequence of the pilot study. Short stories were read to the control group followed by discussions of social studies problems, community helpers, or social roles. In this way, the control group was given special treatment, presumably different from that of the training group only in content.[9]

Instructional Philosophy

The immediate instructional aim of the training study was to teach children to master those structures of operations which characterized the tasks in the classification sequence or the seriation sequence in the pilot study. That is, the training group children were taught to perform tasks which were characterized by the following four groupings of operations: classification, multiplicative classification, seriation, and multiplicative seriation. The criterion used to test the mastery of the above groupings

---

[8]Seventeen children in the pilot study had responded correctly to tasks S1, N2, and C1, but not to tasks S3, N3, or C3. None of these children responded correctly to any of the more difficult tasks in the three sequences.

[9]For a complete account of this training study, including a complete day-to-day account of the training sessions, see Lesh (1971).

of operations was that the child be able to correctly respond to all of the tasks in the classification sequence and the seriation sequence from the pilot study.

To teach the children to master the above four groupings, the following instructional philosophy was employed. For each task which had appeared in either the seriation sequence or the classification sequence of the pilot study, other tasks were devised which were characterized by the same structure of operations but which utilized perceptually different materials. For example, in the case of Task S5 (putting ten dowel rods in order according to length), a set of isomorphic tasks was given in the section of this paper titled "The Genetic Construction of Operational Structures."

Isomorphic sets of tasks relative to each of the tasks which were involved in the classification and seriation sequences of the pilot study were presented following the sequential order of difficulty which had been revealed in the pilot study. Thus, isomorphic sets of tasks were presented first for Task S1, then for C1, then for S2, then for C2, etc.

In the process of encouraging children to actively apply seriation and classification schemes in tasks which required progressively higher degrees of coordination and flexibility, the following training variables were introduced as by-products. Children were taken from tasks which required only a semi-anticipation of the result of the application of an operation to a set of objects, through a trial and error period in which semi-anticipation was successively corrected by hindsight, to a period characterized by the attainment of reversibility of the relevant operations in which the results of operations could be genuinely anticipated (i.e., foresight). Further, children were gradually required to overcome their cognitive tendencies toward centering and syncretic thinking. Nonetheless, it should be stressed that decentration and analytic thinking were considered to be primarily by-products of the gradual coordination of the schemes of actions rather than conversely. The primary focus of the training procedure was to coordinate schemes of actions which lead to the groupings of seriation, multiple seriation, classification and multiple classification.

In addition, children in the training group became acquainted with the meaning of the following words relative to the correct completion of their seriation and classification tasks: *alike, different, order, some, all,* and *more than*. Although the use of language was considered to be useful in helping children to organize their seriation and classification activities, the emphasis of the training procedure was on coordination rather than on the use of language. For the most part, the use of language was left to the spontaneous application by the children. Little effort was made to refine the children's use of language except in the instances noted above.

Posttest: Results

The posttest was given individually to each of the twenty children who participated in the study. The results were striking. None of the ten children who had been in the control group were able to respond correctly to any of the six tasks on the posttest. In contrast, four of the ten children in the training group responded correctly to all of the number tasks (i.e., N5, N6a, and N6b). One other training group child responded correctly to only Task N5b, and one responded correctly to only Task N6b. Thus, the training group significantly out-performed the control group. Of equal importance, however, is the fact that none of the ten children in the training group were able to respond correctly to either of the transformation tasks T6 or T7.

Conclusions

The training group's sessions aimed at teaching children those operations and operational structures which were involved in the seriation and classification tasks of the pilot study. The posttest revealed that transfer of learning from the experience which the training group had received induced improved understanding concerning number concepts. Further, the nature of this transfer was determined by the fact that whereas transfer was obtained to tasks involving number concepts, no transfer was obtained to tasks involving only spatial transformations.

Children who had simply learned to make two simultaneous judgments would have been expected to perform better on not only the number tasks, but also on the transformation task of the posttest. However, this was not the case, even though the number tasks and the transformation tasks had been shown to be comparable in difficulty. Other conjectures such as maintaining that the training group had simply had more practice in dealing with concrete materials must similarly be rejected since one would expect to find comparable improvement on both the number tasks and the transformation tasks of the training study.

The results of this study seem to bear witness to the fact that children may indeed be capable of abstracting the operational essence from the series of tasks which are characterized by isomorphic operational structures. That is, children appear to be capable of internalizing *schemes* of actions. Further, once these schemes are internalized as reversible structures (i.e., groupings), they appear to generalize to new situations which involve the same scheme. Therefore, the results of the study appear to significantly strengthen the psychological viability of ordering and equating operational concepts (or tasks) on the basis of their underlying operational structure. Further, the fact that no transfer was obtained from the seriation and classification instruction to tasks which involve only spatial transformations indicates that Piaget's distinction between three basic types of cognitive logical-mathematical operations may be quite a fruitful sort to make. There may be certain qualitative psychological differences between the types of operations which have been referred to in the present article as operations, relations, and transformations.

Since Piaget's analysis of operational concepts (and number concepts in particular) relies heavily on the psychological viability of analyzing, ordering, and equating tasks on the basis of their underlying operational structures, his description of the development of such concepts was strengthened by the results of this study. For Piaget, operational concepts such as the concept of number need not be assumed to evolve out of some sort of initially nonoperational (e.g., purely perceptual-linguistic) concept. Rather, an operation is an internalized *scheme* of actions which is reversible and which depends on other operations, with which it forms a structured whole characterized by laws of totality (Beth & Piaget, 1966, p. 234).

Two facts should perhaps be mentioned concerning the fact that children who spontaneously mastered all of the tasks in the seriation and classification sequences of the pilot study out-performed the ten children in the training group of the present study.

1. While a concept of number might involve only seriation and classification operations, conservation of number tasks involve transformations as well. For this reason, if the goal is to teach children that number remains invariant under simple spatial transformations, it seems likely that the child will have to be taught all three types of logical-mathematical operations (i.e., operations, relations, and transformations). However, the purpose of this study was not to isolate sufficient conditions which will insure the child's mastery of conservation tasks. Rather, it was to produce transfer of learning due to similarities in underlying operational structures.

2. Cognitive operations are perhaps never mastered in any absolute sense during the period of concrete operations. This is because children are not actually aware of the systems of operations which they bring to bear in various logical-mathematical settings. Mastery of a given operational structure means that a child is able to apply the system of operations in a wide range of concrete situations. Further, it seems likely that the more situations in which a child has learned to apply a given operational structure, the greater will be the chance that the structure can be applied to a new setting. What we have here is near tautology; the wider the applicability of a given cognitive operational structure, the greater will be the chances that the structure will be applied in any given situation. Maximum transfer can be expected from learning of greater generality.

Children in the training group were given a relatively small number of tasks representing each operational structure which characterized the tasks in the seriation and classification sequences of the pilot study. Therefore, it is not unreasonable to suppose that their level of mastery of the relevant structures of operations would have been somewhat less than that of a child who could spontaneously respond correctly to all of the seriation and classification tasks. Thus, the degree of transferability which one would expect from such learning would be somewhat less than for spontaneous problem solvers.

Since training studies are frequently interpreted as encouraging the acceleration of a child's cognitive development, a disclaimer concerning the unqualified desirability of acceleration should be emphasized. From the point of view of this paper, Piaget's theory can perhaps be most useful in providing guidance to broaden the conceptual basis underlying those mathematical topics that are of greatest importance to mathematics educators.

If Piaget's analysis of the development of logical-mathematical concepts is taken seriously, then for many concepts that are most fundamental, learning may have to be much more broadly based than many educators have been willing to admit. So, while acceleration of a single isolated concept is no doubt possible, even this acceleration should only be possible within limits that are imposed by the breadth of the child's conceptual bases.

In order to come to an understanding of the concept of number, children may have to have certain experiences seriating and classifying objects, as well as certain experiences concerning spatial transformations. The exact nature of these experiences may be able to be determined by analyzing, ordering, and equating tasks on the basis of their underlying operational structures.

Figural models can be very useful in order to help children come to an understanding of logical-mathematical concepts. But, according to Piaget before a model can be used as an image to represent a mathematical concept certain systems of operations will usually have to be coordinated relative to the model. For mathematical concepts, figural models are typically constructed representations. That is, mathematical properties will usually have to be put into the object using systems of operations before the properties can be abstracted.

DAVID C. JOHNSON

Learning of Selected Parts of a Boolean Algebra by Young Children

Literature pertaining to theoretical as well as empirical study of the thinking of young children is quite abundant. Although used extensively by psychologists, this literature remains largely untapped by mathematics educators. However, for mathematics educators interested in cognition, the research literature surrounding the work of the Geneva School provides a framework for (1) explaining how mental operations basic to mathematical thought develop, (2) identifying structural characteristics of thought as they undergo change with age, and (3) forming a theoretical basis for certain curricular decisions and experiments in the learning of mathematics.

The present study was designed with the following purposes: (1) To determine if specific instructional conditions improve the ability of young children of various ages and intellectual levels to (a) form classes, (b) establish selected equivalence or order relations; and (2) to investigate that if specific instructional conditions improve abilities outlined in (a) and (b) of (1) above, whether transfer occurs to (a) other class-related activities and (b) the transitive property of the selected equivalence and order relations.

Grouping Structures

Piaget (Beth and Piaget, 1966, pp. 158-162) has identified four main stages in which structural characteristics of thought are qualitatively different. They are: (1) sensory-motor, preverbal stage; (2) the stage of preoperational representation; (3) the stage of concrete operations;

The experiment reported in this chapter is based on a doctoral dissertation in the Department of Mathematics Education at the University of Georgia (Johnson, D.C., 1971).

and (4) the stage of formal operations. Concrete operations are a part of the cognitive structure of children from about 7-8 years of age to 11-12 years of age. Piaget (1964c, p. 42) postulates that this cognitive structure has the form of what he calls groupings, of which five properties exist. Eight major groupings are identified, each of which satisfies the five properties. The idea of an operation is central to these groupings. Piaget (1964c) views an operation as being an interiorized action, always linked to other operations and part of a total structure. Piaget's claim is that operations are fundamental to the understanding of the development of knowledge. The groupings are the structures of which the operations are a part. The difference in the groupings resides in the various operations which are structured. The elements of two groupings are classes and asymmetrical relations which correspond to the cognitive operations of combining individuals in classes and assembling the asymmetrical relations which express differences in the individuals.

It must be made clear that the Geneva School is concerned with describing transformations that intervene between the input of a problem and the output of a solution of the problem by a subject. As Bruner (1959) put it, "Piaget proposes to describe them [the transformations] in terms of their correspondence to formal logical structures [p. 364]." At a certain stage, a child becomes capable of solving a variety of problems not possible at an earlier stage, but is still not able to solve other problems which contain elements of a more advanced stage. In short, Piaget has provided a structure of intelligence which can be used to account for success or failure of children when solving certain problems.

Because the grouping structure is used as a tool to characterize the thinking of the young child, it is interesting to give an interpretation. In *The Psychology of Intelligence,* Piaget apparently selects special classes for part of his elements in the first grouping. These classes must satisfy the following pattern: $\phi \subset A_1 \subset A_2 \ . \ . \ . \subset \cup A_\sigma$, where $\sigma \varepsilon A$ and A is the index set. If "$\subset$" is interpreted to mean "$\subseteq$," then the above sets constitute a lattice, which is a partially ordered system in which any two elements have a greatest lower bound and a least upper bound. Clearly, "$\subset$" is a partial ordering of the sets in question since it is (a) reflexive, (b) antisymmetric, and (c) transitive. Moreover, for any two elements A_α and A_β, $A_\alpha \cap A_\beta$ is the greatest lower bound and $A_\alpha \cup A_\beta$ is the least upper bound.

This lattice structure is not all that is included in the first grouping. Classes of the form $A'_\sigma = A_\gamma - A_\sigma$ where $A_\sigma \subset A_\gamma$ are also included. The classes A'_σ included along with the elements of the lattice are the elements of this first grouping. If one interprets Piaget's (1964c) "+" to be "$\cup$," then he gives (embedded in a zoological classification) statements analogous to the following [p. 42]:

1. Combinativity, $A_\sigma \cup A'_\sigma = A_\gamma$;
2. Reversibility, If $A_\sigma \cup A'_\sigma = A_\gamma$ then $A_\sigma = A_\gamma - A'_\sigma$;
3. Associativity, $(A_\sigma \cup A'_\sigma) \cup A'_\gamma = A_\sigma \cup (A'_\sigma \cup A'_\sigma)$;
4. General Operation of Identity, $A_\sigma \cup \phi = A_\sigma$;
5. Special Identities, (a) $A_\sigma \cup A_\sigma = A_\sigma$, (b) $A_\sigma \cup A_\gamma = A_\gamma$ where $A_\sigma \subset A_\gamma$.

When considering definitions of a Boolean Algebra such as recorded in

Modern Algebra by Birkhoff and Maclane [1958, pp. 336, 337], it can be noted that aspects of a Boolean Algebra are inherent in Grouping I. For example, there are two binary operations "$\cap$" and "$\cup$" with all of the usual properties (such as commutativity and associativity) and a binary relation "$\subseteq$" which orders the subclasses. Also, $\phi \cup X = X$ if X is a class in the system.

Grouping I also describes essential operations and relations involved in cognition of simple hierarchies of classes. Proficiency with the use of the class inclusion relation is viewed by Piaget as essential in the establishment of operatory classification. Two abilities, described by structural properties, are of particular importance in this proficiency. The first is the ability to compose classes (combinativity) and decompose classes (reversibility), and the second is the ability to hold in mind a total class and its subclasses at the same time, made possible through combinativity and reversibility; or as will be seen later, through an ability to think of two attributes at the same time.

Due to the centrality of the class inclusion problem as a test of operatory classification, Piaget (1952) reported an early study with children of ages four to eight. A major part of the investigation involved presenting the children individually with materials similar to the following: wooden beads, the majority of which were brown; blue beads, the majority of which were square; and flowers, the majority of which were poppies. Typical kinds of questions asked were the following: (a) Are there more wooden beads or more brown beads? (b) Would a necklace made of the wooden or of the brown beads be longer? or (c) Would the bunch of flowers or the bunch of poppies be bigger? The questions were quite difficult for children under seven, but children over seven performed quite well. The main reason attributed to the failure of the younger children was that they supposedly could not think simultaneously of the whole and its parts, as mentioned above.

Continuing the "additive" operations, Piaget delineates two groupings entitled "Addition of Asymmetrical Relations" and "Addition of Symmetrical Relations." The asymmetrical relations referred to are interpreted here as strict partial orderings, i.e., orderings that are (1) transitive, (2) asymmetric, and (3) nonreflexive. Moreover, if such relations are linear, then the set A on which the relation is defined is a chain and hence is a lattice. The general properties of a grouping may be applied. Combinativity can be interpreted under the more general notion of relation composition. That is, $A \alpha B$ and $B \alpha C$ implies $A \alpha C$ which is an expression of transitivity. Reversibility by reciprocity includes permuting the terms of the relation as well as reversing the relation, i.e., the reciprocal of $A \alpha B$ is $B \propto A$. The composition is associative by virtue of the transitive property and has special identities. Addition of symmetrical relations involves several distinct categories of relations; some transitive, some intransitive, some reflexive, and some nonreflexive.

Piaget (1964c) also describes groupings based on multiplicative operations, i.e., those which deal with more than one system of classes or relations at a time. Two of these groupings are called Bi-Univocal Multiplication of Classes and Bi-Univocal Multiplication of Relations. In the former, an example is given by the following: If C_1 and C_2 denote the same set of, say, squares, but $C_1 = A_1 \cup A_2$ and $C_2 = B_1 \cup B_2$ where A_1 denotes red squares, A_2 blue squares, B_1 large squares, and B_2 small squares, then $C = C_1 \cap C_2 = (A_1 \cap B_1) \cup (A_1 \cap B_2) \cup (A_2 \cap B_1) \cup (A_2 \cap B_2)$. In other words, a matrix or double entry table of four cells has been generated with the component classes of C_1 on one dimension and those of C_2 on the other. In the case of Bi-Univocal Multiplication of Relations, an example could be seriating a collection of sticks according to lengths and diameter (thickness). A double entry table would thus be defined. If L denotes length and T thickness, then the matrix could look as follows. All the objects in the first row are the same thickness but different lengths while the objects of the first column are the same length but different

L_1T_1	L_2T_1	L_3T_1	L_4T_1	. . .
L_1T_2	L_2L_2	L_3T_2	L_4T_2	. . .
L_1T_3	L_2T_3	L_3T_3	L_4T_3	. . .
L_1T_4	L_2T_4	L_3T_4	L_4L_4	. . .
. . .				

thickness. It must be pointed out, however, that L_1T_1 denotes *at least* zero objects, so that equivalence as well as order relations are potentially involved in this process. The structural properties of these latter two groupings are not discussed--except to say that multiplication of classes allows a child to classify according to two or more classification systems at once--or to consider an object as possessing two or more attributes simultaneously, and that multiplication of relations allows a child to seriate a collection of objects according to two or more order relations at the same time.

In general, classification (which involves equivalence relations) and seriation (which involves asymmetric relations) are at the heart of the theory of Piaget. When asked to classify, children below the age of five usually form "figural collections." By age seven, children can sort objects, add classes (form unions), and multiply classes (cross classify). However, genuine operatory classification does not exist until age eight when children can solve the class inclusion problem. Although $(A + A' = B)$ is logically equivalent to $(A = B - A')$, many children have difficulty with the latter having mastered the former as shown by a failure to state $B > A$ (B contains more than A). The conservation of the whole (being able to hold the class B in mind when focussing on A) and the quantitative comparison of whole and part $(B > A)$ are the two essential characteristics of genuine class inclusion (Piaget, 1964c, p. 117).

Recognizing that empirical research exists which provides evidence for existence of the above groupings (i.e., replications studies) and that experiments exist which have been designed to test the theory (i.e., training studies), the present study was of a slightly different nature, but was embedded in existing psychological, mathematical, and logical theories and structures. Just how it was embedded is made clear as the study is laid out. It must be emphasized that the study was not done to test Piaget's theory or to replicate already known results, such as those produced by Smedslund (1963c), Bruner and Kenney (1966), and Shantz (1967), but an employment of the theory in an applied research problem. To be sure, controversies exist concerning the validity of the theory (e.g., see Kohnstamm (1967), Braine (1959)).

Method

The theory of Piaget is a theory of development which subordinates learning to development in contrast with behavioral theories which attempt to explain development in terms of learning (e.g., Gagné's work). As a corollary, one could view mathematical experiences (e.g., school instruction) as not being assimilated in any genuine way in the absence of requisite cognitive structure. More specifically, it would appear that work on classifications and relations would bear little fruit for children in the stage of preoperational representation. However, as Sullivan (1967) comments: "If learning should be geared to the child's present developmental level as Piaget insists, then the problem of matching the subject matter to the growing conceptual ability of the child (i.e., present cognitive structure) is a relevant consideration [p. 19]."

Learning Material

Classifications and relations were the broad topics about which learning material was constructed. The basic connectives considered in the learning material were conjunction, disjunction, and negation, as well as selected mathematical relations. The learning material, described in detail elsewhere (Johnson, D., 1971), was conducted to provide children with experiences in forming (1) classes, (2) intersection and union of classes, (3) the complement of a class, and (4) relations between classes and between class elements. Physical objects were employed so that each child could be actively involved. Some free play was permitted and interaction with peers was encouraged. The learning material was administered in 17 instructional sessions each lasting about 20 minutes. The first three sessions were designed to provide experiences in forming classes. Hula hoops and other representations of closed curves were used in all sessions to motivate formation of classes. In sessions IV, V, and VI work was done on the intersection and the complement of the intersection of classes. The children were put in a conflict situation when it was pointed out that an object could not be placed inside two nonoverlapping hula hoops simultaneously. For example, if the children were instructed to place red objects in one hula hoop and triangular shapes into another hula hoop, the problem would arise as to where the red triangles should go. Sessions VII and VIII included activities concerning formation of the union of classes. Sessions XII, XIII, XIV, and XV contained activities designed to operationally define the relations "more than," "fewer than," and "as many as." The remaining sessions involved review on formation of classes involving complementation, intersection, and union. Five basic posttests were then constructed to measure achievement and transfer.

Posttests

Connective Achievement Test (CA). The connective test was designed to measure an ability to use the logical connectives "and," "or," and "not." Two sets of physical objects were used in the testing. One set consisted of Dienes Logic Blocks used in the learning sessions (CA_1) and the other set consisted of physical objects which had not been used in the learning sessions (CA_2). Ten items were written using each set where the items were isomorphic across sets except for the differences in the objects used. Six warm-up questions were included for each set of objects to insure that the children understood basic attributes of the objects. The phraseology "Put in the ring *all* the things that are ..." preceded the directions in each of the 12 warm-up and 20 test items. The directions for the ten items involving physical objects which had not been used in the learning sessions were:

1. Either sticks or they are clothespins.
2. Either sticks or they are not blue.
3. Not blue discs.
4. Red discs.
5. Clothespins and they are blue.
6. Either sticks or they are green.
7. Not blue and they are not clothespins.
8. Not red.
9. Discs and they are sticks.
10. Red and they are not sticks.

Relation Achievement Test (RA). This 25 question test was designed to measure understanding of the relations "more than," "fewer than," "as many as," "same shape as," and "same color as." For each of the first three relations, objects used in the items were mounted on pieces of posterboard in a vertical, horizontal, and circular arrangement, for a total of nine items. The set of number pairs used for the "as many as" relation was {(6, 6), (7, 7), (8, 8)}. The set used for the "more than," and "fewer than" relation was {(5, 6), (6, 7), (7, 8)}. A "more than," "fewer than," and "as many as" question was asked for each item to insure that when a child said, for example, "There are more A's than B's," he also knew that there were neither fewer A's than B's nor as many A's as B's. An example question would be, "Are there fewer A's than B's?" For the 16 shape and color items, eight cards (containing two objects each) were constructed, two for each pair in the set {(same shape, same color), (same shape, different color), (different shape, same color), (different shape, different color)}. Each card was used for two items, a shape item and a color item. The tester pointed to the appropriate object and asked: "Is this the same shape as that?" and "Is this the same color as that?" The next three tests to be described are transfer tests with the exception of the intersecting ring items in the Multiplication of Classes and Relations Test.

Multiplication of Classes and Relations Test (MU). This test was constructed to measure the ability of children to use two or more criteria at once. Parts of this test were similar to the nine matrix tasks designed by Inhelder and Piaget (1964, pp. 60-61), which were either four-cell or six-cell matrices with from five to eight choices located below the matrix. For the purpose of testing the ability of children to multiply classes and relations, six material sets spanning across each of the following three types of arrays were utilized: (1) 3 x 3 matrices, (2) 2 x 2 matrices, and (3) ring intersection. The six sets were defined by the pairs in the following set: { (shape used in learning material, color used in learning material), (shape, color used in learning material), (color, number), (shape, shading), (shape, size), (color, size)}. Exactly one material set was used in the construction of each of the six 3 x 3 matrices, of each of the six 2 x 2 matrices, and of each of the six intersecting ring patterns. Although the intersection ring activity was not performed during the unit, it was very similar to some activities and was thus considered as an achievement measure. The matrix items were never solved in the instructional unit and hence were viewed as transfer measures. For each of the eighteen items described, a strip of four response choices was constructed. For the matrix items each response strip included (1) the correct missing object, (2) an object from the same column but a different row than the missing objects, (3) an object from the same row but a different column than the missing object, and (4) an object having one attribute not represented in the matrix. For each pair of intersecting rings, corresponding response strips included (1) an object from the left ring, (2) an object from the right ring, (3) the object logically belonging in both rings (possessing both attributes), and (4) an object having one attribute not represented in either ring.

Class Inclusion Test (CI). This 16 item test was included as a transfer measure for two reasons. First, whenever a class and its complement are specified, the idea of inclusion is implicit. Second, as already noted, Piaget views successful solution of the class inclusion problem as indicative of operatory classification.

Factors affecting the ability to solve inclusion problems are: (1) presence of an extraneous object, (2) three or more proper subsets present, (3) equal numbers in a set and its complement, (4) mingled items, (5) items not visually present, (6) addition or subtraction of an item after initial comparison. These factors were utilized in designing the items of this test. Two other factors included were: (7) items of an infinite nature, and (8) items where subset comparison is made through the use of an outside set of objects. Eight items involved factor 1; two, factor 2; three, factor 3; nine, factor 4; one, factor 5; two, factor 6; two, factor 7; and two, factor 8. With the exception of two items, the number of objects in subclasses for each item was assigned to the items randomly where the numbers were members of the set { 2, 3, 4, 5}.

The first 14 items (items numbered 1-14) included two questions. An example item is where there were five blue tops, three blue guns, and two turkeys. The experimenter had the child point to the toys and to the tops. The two standard questions, "Are there more toys than tops?", and "Are there more tops than toys?" were then asked. The last two items were analogous to one another in that each included Factor 8. In addition to two standard questions, two other questions were asked concerning comparison of the outside set of objects with the set of objects of direct concern. For example, one item contained pictures of seven animals (four horses and three rabbits) and four dots arranged proximal to the horses. Two questions were asked requiring the child to compare the horses and dots, and the animals and dots as well as the two standard questions.

Transitivity Test (TR). This 10 item test was designed to measure the ability of children to use the transitive property of the relations tested for in the Relation Achievement Test. Two items were designed to test for the transitivity property of each of the five relations. A "left to right" and a "right to left" matching were used in the testing for the transitive property of the relations "as many as," "more than," and "fewer than." The triplets of numbers of objects used for testing for the above three relations were (7, 7, 7) and (8, 8, 8) ; (8, 7, 6) and (9, 8, 7) ; and (6, 7, 8) and (7, 8, 9) respectively. The test was used as a transfer measure to determine if an ability to use transitivity is improved by instruction on the relations of concern.

An example of a transitivity item for matching relations is where there were seven red discs and seven green discs mounted in rows on posterboard. The child was directed to match a pile of seven blue discs with the red discs and judge the relation between the two sets. The red discs were then covered. The child was then directed to match the blue discs with the green discs and judge the relation between the two sets. The green discs were then covered. Three questions were then asked; "Are there as many red discs as green discs?" "Are there more red discs than green discs?" and "Are there fewer red discs than green discs?". An analogous procedure was used for transitivity of the equivalence relations involving color and shape, except only two questions were asked, one for the appropriate equivalence relation and one for its accompanying difference relation.

Sample

The subjects for the study were chosen from four kindergarten and four first grade classes located in or closely adjacent to Athens, Georgia. All of these children were administered an Otis-Lennon Ability Test during March 24-April 1, 1970. A total of 99 first graders and 97 kindergarteners were tested. Two levels, Primary 1 and Elementary 1, of the Otis-Lennon Mental Abilities were utilized. The Primary 1 level is designed for pupils in the last half of the kindergarten and Elementary 1 level is designed for pupils in the last half of the first grade. The test items sample the mental processes of classification, following directions, qualitative reasoning, comprehension of verbal concepts, and reasoning by analogy. K-R 20's for the Primary and Elementary Levels are .88 and .90 respectively. The two categorization variables, then, were chronological age and IQ. Only those children who had an IQ in the interval (80, 125) and a CA either in the interval (64, 76) or (77, 89) for kindergarten and first grade, respectively, were included in the study. The children were further categorized by the two IQ intervals (80, 100), (105, 125). Children within the four categories thus defined were then randomly assigned to an experimental or control group after an ordered random sample of 80 subjects had been selected, 20 in each category. Thirty-five alternates were also selected for a total of 115 children in the sample.

Administration of the Tests

Administration of CA. The CA was administrated to six subjects at a time. Three subjects were seated adjacent to each other on one side of a table and the other three were seated facing them on the opposite side of the table. Subjects were separated by cardboard partitions so they could not see each other. Each subject was given a rope ring and some objects to classify. No objects were initially inside the rope rings. The order of test questions was initially randomized. The investigator read all the directions clearly and repeated if necessary. All subjects were given sufficient time to make their responses. The experimenter stood behind the subjects and recorded each response as correct (correct set of objects was placed in ring) or incorrect (either items omitted or at least one incorrect item placed in ring).

For subtest CA_1, if all the proper objects were placed in the ring and nothing extra was placed there the answer was considered as correct. One point was given for correct answers and no points were given for incorrect answers. Subtest CA_2 was scored in a similar way. Since the tests were parallel, Subtest CA_3 was formed through the consideration of the responses to the items in Subtests I and II. The subjects were given credit for having a question right on Subtest III only if they had scored each corresponding question right on both CA_1 and CA_2. In considering Subtest CA_3, one point was given for each question judged as right by the above procedure. The normal testing time was approximately 23 minutes.

Administration of RA. For this test, the material sets were placed in a row on a low table in order from 1 to 17. Administration of items 1-9 (matching relations) was done first with the sequence of presentation randomized individually for each subject. Also the question sequence was randomized for each question for each subject. Cards 10-17 (shape and color relations) followed with the sequence of presentation also randomized for each subject. Here again, the question sequence was randomized for each subject. The eight "same shape" questions asked of cards 10-17 composed items 10-17 for this test and the eight "same color" questions composed items 18-25 respectively. For each card, the response was scored correct if the color and shape questions were both correct. The tester recorded the "yes" and "no" responses for each question asked. Average testing time was approximately twelve minutes.

Administration of MU. The eighteen material sets for this test were placed in order (1-18) on a low table similar to that used with the RA. Each strip of four response choices was centered and placed directly below the respective matrix or ring item. The sequence of presentation of the eighteen items was randomized for each subject. The tester recorded the response choice pointed to on each response strip. Average testing time was approximately twelve minutes.

Administration of CI. The 16 items were partially randomized for each subject. The exceptions were that items numbered 3 and 4, and 13 and 14 were presented in pairs in the natural order and items numbered 15 and 16 were presented last in the natural order. An item was scored

Table 1
Formation of Subtests

Test Type	No. of Items	Subtests	Content of Subtest
Achievement Tests	10	CA_1	First ten items of CA
	10	CA_2	Last ten items of CA (novel material)
	10	CA_3	Intersection of Tests CA_1 and CA_2
	25	RA	Same as RA
	6	MU_r	Last six items of MU (intersection rings)
Transfer Tests	6	MU_3	First six items of MU (3x3 matrices)
	6	MU_2	Second six items of MU (2x2 matrices)
	16	CI	Same as CI
	10	TR	Same as TR

as correct only if the two standard questions were correctly answered.

Administration of TR. Items were arranged in a row on a low table. Administration of the six items for matching relations was conducted followed by the four items for the color and shape relations. Within this constraint, the items were randomized independently for each subject. A transitivity item was scored as correct only if all questions were correctly answered.

Design of Study and Method of Analysis

The basic design of the study was The Posttest-Only Control Group Design presented by Campbell and Stanley (1963). This design calls for initial randomization followed by an experimental treatment given to the experimental group. Twelve major multivariate analyses of variance null hypotheses were tested.

H_1: The mean vectors of the experimental and control groups are not different on the achievement measures.
H_2: The mean vectors of the experimental and control groups are not different on the transfer measures.
H_3: The mean vectors of the kindergarten and first-grade subjects are not different on the achievement measures.
H_4: The mean vectors of the kindergarten and first-grade subjects are not different on the transfer measures.
H_5: The mean vectors of the low and high IQ subjects are not different on the achievement measures.
H_6: The mean vectors of the low and high IQ subjects are not different on the transfer measures.
H_7: There is no significant interaction of IQ with Treatment on the achievement measures.

H_8: There is no significant interaction of Grade with Treatment on the achievement measures.
H_9: There is no significant interaction of Grade with IQ on the achievement measures.
H_{10}: There is no significant interaction of IQ with Treatment on the transfer measures.
H_{11}: There is no significant interaction of Grade with Treatment on the transfer measures.
H_{12}: There is no significant interaction of Grade with IQ on the transfer measures.

Test statistics and an item analysis were computed for each of the subtests composing the transfer and achievement measures. Two point biserial correlation coefficients, a phi coefficient, and a difficulty index were computed for each item. A point biserial correlation coefficient represents the degree of correlation existing between a dichotomous and a continuous variable. In the study, IQ measures and the total test score formed by the composite of posttest scores were considered continuous variables. The dichotomous variables are the individual items scored as either correct or incorrect. Correlations involving IQ and total scores provide indices of validity and reliability respectively. Essentially, a phi coefficient is, with minor modification, a chi-square calculated on a two-way contingency table to test for independence of two random variables. The table was defined by experimental and control groups, and the ratio of subjects passing or failing each item to the total responses on that item.

The null hypotheses were tested with the use of Multivariate Analysis of Variance (MANOVA) procedures. Program MUDAID (Multivariate, Univariate, and Discriminant Analysis of Irregular Data) was used for the MANOVAs where the five achievement and four transfer measures were the response variables for all combination of independent variables taken two at a time. Therefore six MANOVAs and 27 ANOVAs were calculated; one for each IQ (I) by Age (A), IQ by Treatment (T), and Age by Treatment. Levels of IQ were 80-100 (L) and 105-125 (H); levels of Treatment were experimental (E) and control (C); and levels of Age were five-year-olds (K) and six-year-olds (F).

Results

The results of the analyses are presented in this section. All data analyzed in the item analysis section were obtained from all 111 subjects and alternates administered all the posttest measures. The multivariate analyses are limited to 80 subjects selected for the study.

Item Analysis

A phi coefficient was calculated for each of the 99 items. Utilizing a significant ϕ ($p < .05$), items which were discriminators between the experimental and control groups were found for each test. From the array of data in Table 2, it can be easily seen that there was only one item which discriminated in favor of the control group out of the total 99 items.

Table 2
Frequency of Items: Discriminators and Nondiscriminators

		Discriminators		Nondiscriminators
No. of Items	Subtest	Experimental	Control	
10	CA_1	8		2
10	CA_2	7		3
10	CA_3	8		2
25	RA	7	1	17
6	MU_r	5		1
6	MU_3	4		2
6	MU_2	2		4
16	CI	0		16
10	TR	7		3

Two of the subtests deserve special discussion in that all or a majority of the items of those tests were nondiscriminators. First, in the case of the RA test, the 16 items which involved usage of the relations "same shape as" and "same color as" were extremely easy for all subjects, and thereby were excluded from all other analyses. Second, four of the six items composing the MU_2 test were nondiscriminators. It appeared that much guessing was done on this test, as the average score was approximately the same as chance would allocate. One of the four nondiscriminators on MU_2 was excluded from all further analyses. Ten other items were also excluded from the analysis with undesirable item characteristics (very hard or very easy items with low or negative biserial correlations with the total test or IQ). Nine of these ten items were nondiscriminators; six for the achievement measures and three for the transfer measures. Seventy-two items were retained for the analysis of variance.

Multivariate and Univariate Analysis

The necessary subtest information is tabulated in Table 3. The internal-consistency reliabilities are quite substantial indicating good homogeneity of the test items. The multivariate and univariate analyses of variance are given for the direct achievement measures (CA_1, CA_2, CA_3, RA, MU_r) and transfer measures (MU_3, MU_2, CI, TR) for the two classification variables (Age and IQ) each considered in conjunction with the treatment variable, and also considered in conjunction with each other.

Table 3
Subtest Statistics

No. of Items	Subtest	Reliability (KR-20)	Grand Mean
9	CA_1	.72	5.09
7	CA_2	.65	3.70
9	CA_3	.74	3.88
9	RA	.82	5.89
5	MU_r	.67	1.39
6	MU_3	.70	3.23
5	MU_2	.58	2.35
13	CI	.75	3.78
9	TR	.79	6.13

Analyses of achievement measures. For the purpose of testing the hypotheses related to achievement, the five achievement subtests were considered concomitantly as response variables in the MANOVA and were considered singly in ANOVAs. In the MANOVA analysis of T vs I, the likelihood ratio test statistic $\chi^2 = 113.30$ was significant ($p < .01$), indicating significant differences in the mean vectors for all effects. As indicated in Table 5, the main effects due to T and I and the interaction of T and I were significant. The test of all F values in Table 5 is done

Table 4
Subclass Means: T vs. I (Achievement Subtests)

Subtest	Low	High	Means
Experimentals			
CA_1	5.35	7.40	6.38
CA_2	4.15	5.95	5.05
CA_3	4.25	6.35	5.30
RA	5.50	8.60	7.05
MU_r	1.55	2.85	2.20
Controls			
CA_1	3.20	4.40	3.80
CA_2	2.05	2.65	2.35
CA_3	1.80	3.10	2.45
RA	3.25	6.20	4.72
MU_r	0.55	0.60	0.58
Means CA_1	4.28	5.90	5.09
CA_2	3.10	4.30	3.70
CA_3	3.02	4.72	3.88
RA	4.38	7.40	5.89
MU_r	1.05	1.72	1.39

Table 5
F Values for MANOVA of Achievement Subtests[a]

Analysis	Factor	F
T vs. I	T	29.66**
	I	10.06**
	T x I	2.52*
T vs. A	T	20.32**
	A	< 1
	T x A	1.13
I vs. A	I	5.43**
	A	< 1
	I x A	1.13

a* = .05 level of significance
** = .01 level of significance

using p and (N-3-p) degrees of freedom where p is the number of response variables and N is the number of subjects. In this analysis p is 5 and N is 80. Also, $F_{.05}$ (5, 72) = 2.35 and $F_{.01}$ (5, 72) = 3.28.

In order to further interpret the main effects of T, I, and T x I, five univariate analyses were performed. The results in terms of F values for these analyses and also for T vs A and I vs A are included within Table 6. It is noted that for each of the five response variables there existed a significant F ($p < .01$) for both T and I. This indicated that performance of children of E and C and also of L and H were significantly different

Table 6
ANOVA F Values for Achievement Measures[a]

Type Variation	CA_1	CA_2	CA_3	RA	MU_r
T	60.22**	114.20**	80.54**	17.31**	44.97**
I	23.98**	22.56**	28.65**	29.30**	7.76**
T x I	1.64	5.64**	1.59	< 1	6.65**
T	13.37**	8.92**	13.96**	23.90**	4.69*
A	< 1	1.25	< 1	< 1	< 1
T x A	< 1	< 1	< 1	< 1	1.09

a* = .05 level of significance
** = .01 level of significance

on all achievement subtests.

Significant interaction ($p < .05$) of T with I occurred only on CA_2 (involving "and," "or," and "not") and MU_r. The significant interaction indicates that on these subtests the performance of control subjects was not like the performance of experimental subjects across the two levels of IQ. Table 4 indicates that on these subtests, the higher IQ experimental subjects performed better than any other group.

In the MANOVA analysis of T vs A, the likelihood ratio test statistic $X^2 = 71.43$ was significant ($p < .01$), indicating significant differences in the mean vectors for all effects. The only main effect that

Table 7

Subclass Means: T vs. G (Achievement Subtests)

Subtest	Kindergarten	First Grade	Means
Experimentals			
CA_1	6.35	6.40	6.38
CA_2	5.00	5.10	5.05
CA_3	5.30	5.30	5.30
RA	6.75	7.35	7.05
MU_r	2.40	2.00	2.20
Controls			
CA_1	3.45	4.15	3.80
CA_2	1.95	2.75	2.35
CA_3	2.00	2.90	2.45
RA	4.90	4.55	4.72
MU_r	0.40	0.75	0.58
Means			
CA_1	4.90	5.28	5.09
CA_2	3.48	3.92	3.70
CA_3	3.65	4.10	3.88
RA	5.83	5.95	5.89
MU_r	1.40	1.38	1.39

was significant in this analysis, as indicated in Table 5, was T. Again, univariate analyses were performed to further interpret the main effect. As shown in Table 6, significance ($p < .01$) was achieved on each of the five subtests if and only if the effect was T.

The final two-way analysis dealt with the factors of I and A. The likelihood ratio test statistic $X^2 = 27.41$ was significant ($p < .01$) indicating significant difference in the mean vectors presented for all effects. As indicated in Table 5, the only main effect that was significant was I. Hence, for the effects of I and A, considered concomitantly, significant

Table 8
Subclass Means: I vs. A (Achievement Subtests)

Subtest	Kindergarten	First Grade	Means
Low			
CA_1	4.00	4.55	4.28
CA_2	2.80	3.40	3.10
CA_3	2.75	3.30	3.02
RA	4.40	4.35	4.38
MU_r	0.90	1.20	1.05
High			
CA_1	5.80	6.00	5.90
CA_2	4.15	4.45	4.30
CA_3	4.55	4.90	4.72
RA	7.25	7.55	7.40
MU_r	1.90	1.55	1.72
Means			
CA_1	4.90	5.28	5.09
CA_2	3.48	3.92	3.70
CA_3	3.65	4.10	3.88
RA	5.82	5.95	5.89
MU_r	1.40	1.38	1.39

differences on achievement existed between the two levels of intelligence used in the study. Table 6 shows that again all F values for the I effect were significant ($p < .01$). As can be seen from Table 8, for all five subtests the mean scores of the high intelligence group were greater than for the low intelligence group and first graders performed better (but not significantly) than or approximately equivalent to kindergarteners. On the basis of the results listed in Tables 5 and 6, hypotheses H_1, H_5, and H_7 were rejected and H_3, H_8, and H_9 were accepted. Hence, for the achievement scores, the factors IQ and Treatment significantly affected performance. First graders performed better, but not significantly better, than kindergarteners on all achievement measures.

Analyses of Transfer Measures

The four transfer subtests were the response variables considered concomitantly in MANOVAs and separately in ANOVAs for the purpose of testing the hypothesea related to transfer effects. For the MANOVA analysis of T vs I, the likelihood ratio test statistic $\chi^2 = 60.19$ was significant ($p < .01$) for all effects. As illustrated in Table 10, the main effects due to T and I were significant but the interaction of T with I was not significant. The test of all F values in Table 10 is done using p and (N-3-p) degrees of freedom as was the case with the achievement

Table 9
Subclass Means: T vs. I (Transfer Subtests)

Subtest	Low	High	Means
Experimentals			
MU_3	3.15	4.45	3.80
MU_2	2.10	3.40	2.75
CI	2.20	5.00	3.60
TR	5.80	8.35	7.08
Controls			
MU_3	2.60	2.70	2.65
MU_2	1.85	2.05	1.95
CI	3.10	4.80	3.95
TR	4.15	6.20	5.18
Means			
MU_3	2.88	3.58	3.22
MU_2	1.98	2.72	2.35
CI	2.65	4.90	3.78
TR	4.98	7.28	6.12

Table 10
F Values for MANOVA of Transfer Subtests[a]

Analysis	Factor	F
T vs. I	T	7.18**
	I	11.75**
	T x I	1.00
T vs. A	T	5.69**
	A	< 1
	T x A	< 1
I vs. A	I	9.68**
	A	< 1
	I x A	< 1

[a]** = significance of factors beyond the .01 level

measures. However, for the transfer measures p is 4 and N is 80. For the new value of p, $F_{.05}$ (4, 73) = 2.49 and $F_{.01}$ (4, 73) = 3.59.

To assist the investigator in interpreting the main effects of T, I, and T x I more precisely, four univariate analyses were performed. F values for these analyses and also T vs A and I vs A are reported in Table 11. For MU_3 and TR significance was maintained ($p < .01$) for the main effect T.

Table 11
ANOVA F Values for Transfer Measures[a]

Type Variation	MU_3	MU_2	CI	TR
T	8.80**	5.59*	< 1	18.95**
I	3.26	4.91*	13.33**	27.77**
T x I	2.40	2.64	< 1	< 1
T	8.25**	5.11*	< 1	14.03**
A	< 1	< 1	< 1	< 1
T x A	< 1	< 1	1.86	< 1
I	2.88	4.45*	13.30**	22.47**
A	< 1	< 1	< 1	1.06
I x A	< 1	< 1	< 1	< 1

a* = .05 level of significance
** = .01 level of significance

A significant F ($p < .05$) was computed for MU_2 but a nonsignificant F was computed for CI. The results were slightly different for the main effect of I. Here, significance ($p < .01$) was established for CI and TR, and for MU_2 there was significance at the .05 level. No significance was found for the main effect of I on MU_3. It is not known why the main effect of I was significant for MU_2 and not for MU_3. One possible explanation is that the subjects of greater intelligence were able to use the fewer cues available in MU_2 more proficiently than subjects of lesser intelligence. Table 9 indicates that significant differential performance always favors the experimental and high IQ groups.

For the MANOVA performed on the pair of factors T and A, the likelihood ratio test statistic $X^2 = 26.04$ was significant ($p < .01$), indicating significant differences for all effects. Only the main effect of T was significant ($p < .01$) as indicated in Table 10. Treatment was significant ($p < .01$) for MU_3 and TR, and was significant ($p < .05$) for MU_2, as given in Table 11. Hence, for those three variables, performance of subjects in the two levels of T differed significantly. Table 12 reveals that for all variables for which the main effect of T was significant, Experimentals outperformed Controls.

The last two-way analysis was done with the pair of factors I and A. The likelihood ratio test statistic $X^2 = 35.48$ was significant ($p < .01$) indicating significant differences in the mean vectors for all effects. As illustrated in Table 10, only the main effect of I was significant ($p < .01$). Table 11 reveals that the main effect of I was significant ($p < .01$) for

Table 12
Subclass Means: T vs. G (Transfer Subtests)

Subtest	Kindergarten	First Grade	Means
Experimentals			
MU_3	3.60	4.00	3.80
MU_2	2.75	2.75	2.75
CI	3.45	3.75	3.60
TR	6.90	7.25	7.08
Controls			
MU_3	2.60	2.70	2.65
MU_2	1.80	2.10	1.95
CI	4.70	3.20	3.95
TR	4.85	5.50	5.18
Means			
MU_3	3.10	3.35	3.22
MU_2	2.28	2.42	2.35
CI	4.08	3.48	3.78
TR	5.88	6.38	6.12

Table 13
Subclass Means: I vs. G (Transfer Subtests)

Subtest	Kindergarten	First Grade	Means
Low			
MU_3	2.60	3.15	2.88
MU_2	1.95	2.00	1.98
CI	3.05	2.25	2.65
TR	4.70	5.25	4.98
High			
MU_3	3.60	3.55	3.58
MU_2	2.60	2.85	2.72
CI	5.10	4.70	4.90
TR	7.05	7.50	7.28
Means			
MU_3	3.10	3.35	3.22
MU_2	2.28	2.42	2.35
CI	4.08	3.48	3.78
TR	5.88	6.38	6.12

CI and TR and was significant ($p < .05$) for MU_2. Hence, IQ plays an important role in performance measured by those variables. No other significant main effects were found. Table 13 indicates that responses favored the high intelligence and first-grade levels.

From the results indicated in Tables 10 and 11, hypotheses H_2 and H_6 were rejected and H_4, H_{10}, H_{11}, and H_{12} were accepted. Therefore transfer to related areas was found to differ significantly depending on levels of I and T. As with the achievement measures, the more intelligent subjects performed better than the less intelligent subjects and the experimental subjects performed better than the control subjects.

Discussion

There is substantial evidence in this study that kindergarten and first-grade children can be taught (1) to form classes using intersection, union, and negation, and (2) to make correct "prenumber" comparisons of sets of objects. Mastery was not required, although significant differences were noted between Experimentals and Controls. Furthermore, this increase in achievement was accompanied by some transfer to related activities. The main effects of Treatment and IQ were very significant on both achievement and transfer measures but the main effect of Age was not significant on any measure.

It is quite important for understanding the results of this study to distinguish between two types of experience--physical experience and logical-mathematical experience. According to Piaget (1964a) physical experience "consists of acting upon objects and drawing some knowledge about the objects by abstraction from the objects [p. 11]." Piaget (1964a) states further that in logical-mathematical experiences "knowledge is not drawn from the objects, but it is drawn from the actions effected on the objects [p. 12]." If a child is asked to place all the objects possessing a given attribute inside a ring, he can be shown his mistakes and they can be corrected. This type of activity is basically in the realm of physical knowledge. However, suppose that a child claims that there are more dogs than animals after he has pointed to the dogs and animals independently. It is impossible to correct his mistakes in a way similar to that of the previous example. With the exception of the MU_r subtest, all the achievement measures fell in the realm of physical knowledge. Hence, the treatment was very effective for imparting physical knowledge. However, the MU_r subtest and the transfer measures must be considered when investigating the production of logical-mathematical knowledge.

Activities with intersecting rings were provided in the unit but in a format that differed from the intersecting ring test items. Although Experimentals performed significantly better than Controls on the MU_r subtest, it can be noted that neither group performed extremely well. Furthermore, Controls appeared to consider the three regions formed by the intersecting rings as nonoverlapping regions. Hence, improvement can be explained by hypotheses other than a genuine improvement in the formation of intersections. In the case of the CI subtest, the treatment did not produce significant differences. On this measure, intelligence produced the only significant effect. However, operatory classification was not achieved by either IQ group because the higher IQ group scored about 37 percent and the lower group only scored about 25 percent, where the expected mean based on guessing is 20 percent. Improvement on the transitivity items can be attributed to clarity of language rather than to usage of the transitivity property. Items based on the relation of shape and color contributed greatly to the rather high mean scores of the Transitivity subtest. Mean scores for Controls and Experiments on matching relations were 30 and 55 percent, respectively, whereas the analogous mean for the shape and color relations were 86 and 97 percent, respectively. The matrix items provided the strongest evidence for an improvement in logical thinking, although the Genevans claim that it is difficult to distinguish between graphic and operational solutions. There was some evidence that the most substantial improvement existed for the high ability first graders.

In conclusion, the unit produced substantial improvement in physical knowledge but very little improvement in operatory classification. When considering the results of the study and observing the way in which addition and subtraction are presented in school mathematics curricula, a serious problem is revealed in that children are being presented with concepts they are conceptually unable to handle. In a subtraction problem such as 9 - 5 = 4, if a child thinks that the difference is larger than the minuend he might just as well write something like 5 - 9 = 4.

Although there was nearly a significant difference in achievement between kindergarten and first-grade children on CA_2, it is recommended that instruction similar to that used in the unit begin at the kindergarten level because there were no significant differences in achievement between these grades on any subtest. However, more research with a more generalized population is highly recommended before final grade-level placement is decided upon. For example, a much deeper investigation is needed concerning the actual relations that exist between the words "and," "or," and "not" and the growth of conjunction, disjunction, and negation concepts respectively. These should be investigated at various grade levels in conjunction with other concepts such as conservation of various relations as discussed by Piaget. The positive transfer made to the transitive property of the equivalence and order relations used in the unit was an interesting outcome. Various properties of the multitude of equivalence and order relations existing in the mathematics curriculum warrant similar investigations. It was noted that relations such as "same shape as" and "same color as" and the transitive property of these relations were very easy even for kindergarteners. Very little, if any, instruction is required in kindergarten for such relations.

IQ should be considered when arranging instruction based on the concepts in this study. Three of the reasons for this are as follows: (1) there was significant interaction ($p < .05$) of treatment with IQ on MU_r with the best performance by the high IQ subjects, (2) among the best discriminators between levels of intelligence was RA, and (3) the intelligence factor was significant on the transfer subtests CI and TR. This is worthy of note because these two subtests occupy key positions in the theory of Piaget. IQ was the only factor where significance was attained for CI. In such areas as those just mentioned, a thorough analysis needs to be made concerning the relation that exists between Piaget's classification of mental operations and the degree to which these operations are measured on various IQ tests. Such investigation could have far-reaching implications for arranging mathematics instruction at various age levels.

At this point in time it is uncertain exactly what abilities the 3 x 3 and 2 x 2 matrix questions and the intersecting ring questions are measuring. There exists good, but inconclusive, evidence that the intersecting ring questions are measuring the same type of ability as the matrix questions. Future investigations need to incorporate other methods when investigating the intersection concept. It is assumed that the improvement in cross classification was done through the "intersection of attribute" activities of the unit. However, it is strongly recommended that the relation existing between two attributes and a total cross classification be investigated further. As indicated previously, Piaget has hypothesized that cross classification, as measured by matrix activities, develops at about age seven and the intersection of simple attributes at about age nine. The present study shows that instruction in one area will perhaps hasten the development of the other operation. Any such transfer is important to education.

THOMAS P. CARPENTER

The Performance of First- and Second-Grade Children on Liquid Conservation and Measurement Problems Employing Equivalence and Order Relations

Conservation or invariance of a given property under certain transformation is basic to the process of measurement of that property. One of the essential features of measurement or comparison of quantities is that the transformations used in the measurement or comparison process do not change the relation between the quantities.

Studies by Piaget, Inhelder, and Szeminska (1960) indicate that this logical interdependence of conservation and measurement is reflected in the development of these concepts in children. However, although Piaget et al. (1960) extensively document relationships between conservation and measurement failures in a variety of situations, their tasks share certain common features which may have influenced their conclusions. First, most of the measurement tasks required relatively sophisticated measurement manipulations. Second, in all comparisons distracting cues were perceptual; and if correctly applied, measurement processes yielded the correct response. There is evidence that certain conclusions of Piaget et al. (1960) resulted from this lack of experimental variability.

They concluded that young children are dominated by the immediate perceptual qualities of a situation. However, the results of another investigation (Carpenter, 1971a) indicate that young children respond to numerical cues with about the same degree of frequency as perceptual cues. As a consequence the children in this study demonstrated an ability to interpret and apply aspects of the measurement process earlier than indicated by Piaget et al. (1960).

In the current study the relation between young children's responses to perceptual and numerical cues on liquid conservation and measurement problems was systematically investigated. A second dimension of the study was to investigate the effect of equivalence and order relations on young children's performance on conservation and measurement of liquid quantities.

The experiment reported in this chapter is based on a doctoral dissertation in the College of Education at the University of Wisconsin (Carpenter, T. P., 1971).

A third dimension of the study was to provide insight into young children's conception of a unit of measure and their understanding of the relation between unit size and number of units.

Mathematical Definition of Measurement

Mathematically, measurement can be discussed in terms of a function mapping the elements of a given domain into some mathematical structure (usually a subset of the real numbers) in such a way as to preserve the essential characteristics of the domain. First a structure must be established on the domain by applying some empirical procedures to define equivalence and order relations ($\sim$ and $<$ respectively) to compare elements of the domain. Generally order relations are established by demonstrating equivalence between one quantity and a proper subset of the other quantity. Thus, logically the definition of equivalence relations precedes the definition of order relations. Similarly empirical procedures are used to define an operation "$*$" that is both commutative and associative to combine elements of the domain.

Specifically, in the case of liquid measure we could say that two quantities of liquid are equivalent if they both exactly fill identical containers. Quantity A is greater than quantity B if it fills one of the containers with some left over. The operation is defined by simply pouring one quantity of liquid into the other.

Once the domain has been given a recognizable structure, a function μ that maps the domain into a subset of the real numbers and preserves the essential characteristics of the structure of the domain must be defined. For liquid measurement this means that given liquid quantities ℓ_1, ℓ_2, and ℓ_3

1. $\mu(\ell_1) = \mu(\ell_2)$ if and only if $\ell_1 \sim \ell_2$

2. $\mu(\ell_1) < \mu(\ell_2)$ if and only if $\ell_1 < \ell_2$

3. $\ell_3 \sim \ell_1 * \ell_2$ implies that $\mu(\ell_3) = \mu(\ell_1) + \mu(\ell_2)$ assuming that ℓ_1 and ℓ_2 do not intersect.

The measurement function for liquid quantities is defined by arbitrarily selecting a quantity of liquid ℓ_0 as a unit. Then any other quantity of liquid is compared with successive multiples of ℓ_0 until a multiple $n\ell_0$ is found such that $n\ell_0$ is less than or equivalent to the given quantity which in turn is less than $(n + 1)\ell_0$. (A multiple $n\ell_0$ is defined to be a quantity of liquid equivalent to $\ell_0 * \ell_0 * \ldots * \ell_0$ in which there are n terms.) Next a quantity ℓ_1 is chosen such that $10\ell_1$ is equivalent to ℓ_0, and a multiple of ℓ_1 is joined to $n\ell_0$ such that $n\ell_0 * n_1\ell_1$ is less than or equivalent to the given quantity which in turn is less than $n\ell_0 * (n_1 + 1)\ell_1$. Similarly ℓ_2 and n_2 are chosen such that $n\ell_0 * n_1\ell_1 * n_2\ell_2$ is less than or equivalent to the given quantity which is less than $n\ell_0 * n_1\ell_1 * (n_2 + 1)\ell_2$. Continuing in this manner a decimal number $r = n.n_1n_2n_3\ldots$ can be constructed and used to define the function μ mapping the domain of liquid quantities onto the set of positive real numbers by $\mu(\ell) = r$, where ℓ is the given quantity above.

When the function μ_0 is defined by arbitrarily selecting a quantity ℓ_0 as a unit, a different function μ_1 can be defined by choosing a different quantity ℓ_1 and using it to generate μ_1. For liquid measurement functions the relation between the functions μ_0 and μ_1 is the form

$\mu_0 = k\mu_1$, where k is a positive real number and $k > 1$ when $\ell_0 < \ell_1$ and $0 < k < 1$ when $\ell_0 > \ell_1$. In other words, for these functions there is an inverse relationship between the unit size and the number of units.

It should be noted that a basic assumption in attributing a structure to the domain and defining the function from the domain to the set of positive real numbers is that the property that is being measured does not change under certain transformations and is not affected by the empirical procedures used to define the relations and operation on the domain. This assumption pervades the entire measurement process. One of the essential characteristics of a measurement function is that it preserves the relation between elements of the domain that it measures. Thus, it is critical that neither the empirical procedures employed to compare elements of the domain directly or the procedures used to define and apply the measurement function affect the relation between elements.

Related Research

For Piaget et al. (1960) this assumption that certain properties remain constant under certain transformations is the central idea underlying all of measurement. It is upon this assumption that these authors have based their investigations of the development of measurement concepts. Based on their studies of length, area, and volume they proposed a stagewise development of measurement which is interrelated with the development of conservation.

The measurement problems in these studies can be divided into two broad classes which correspond to the two major divisions within the mathematical definition of measurement described above, those strictly employing empirical procedures directly to elements of the domain and those employing the measurement function. In the first class of problems, objects of the domain were directly compared on the basis of a given attribute without assigning a number to the attribute. These problems include the classical conservation problems in which one of two objects equivalent in some way is transformed to appear larger or smaller than the other. To conserve, a child must not respond on the basis of the immediate appearance of the objects but rather must recognize that the objects were equivalent in the earlier state and the relation between them did not change. All comparisons were visual and measurement functions were not introduced. For example, in studying the development of area concepts, Piaget asked children to compare two identical rectangles made up of six squares each and arranged in a two by three configuration. After a child agreed that the two rectangles were the same size, the squares in one of the rectangles were moved to create a different shaped region. The child was then asked to compare the size of this new region with that of the undistorted rectangle.

In the second class of problems, measurement functions were applied in which different units of measurement were used. In some problems children were given several different sizes of units and in others the units were such that fractional parts were required to cover the object being measured. In this class of problems the objects being compared were never visually comparable. Thereby, in order to accurately judge the correct relation measurement was necessary. Furthermore, children were required to perform the measurement operations so that they not only had to correctly apply the information from the measurement process, but also had to carry out the measurement manipulations. For example, in studying the development of area measurement functions, children were asked to compare different shaped figures by measuring with different units. In one set of problems children were given enough units to cover the figures, but the units were of different sizes and shapes. Some were squares, some were rectangles (two squares), and some were triangles (squares cut diagonally in half). In another set of problems children were given a limited number of square cards which they had to move by successive iteration from one part of the region being measured to the next. Some regions were shaped in such a way that it was impossible to cover them with the given units without intersecting the exterior of the region. Thus, it was necessary for the children to consider fractions of units.

Based on children's responses to these problems Piaget et al. (1960) concluded that the development of measurement and conservation is integrally related and that the same general pattern of development persists across all types of measurement operations. The earliest stages (Stages I and IIA) are characterized by a dependence on one dimensional *perceptual* judgments. Conservation is not present and transformations from prior stages are completely ignored. Children are unable to apply measurement processes in any meaningful manner; and quantities are compared on the basis of a single, immediate, dominant dimension. In Stage IIB children begin to make a number of correct judgments as long as distortions in quantities being compared are not too great. Correct judgments in Stage IIB are largely a result of trial and error. Children have a dim concept of conservation and some notion that greater quantities measure more units. In Stage IIIA children begin to conserve and measure using a common unit of measure. However, they fail to recognize the importance of a constant unit of measure and often count a fraction of a unit as a whole or equate two quantities that measure the same number of units with different size units of measure. In Stage IIIB children successfully conserve and measure. They recognize the importance of different units of measure and understand the inverse relationship between unit size and number of units. It is not until Stage IV, however, that children finally discover the mathematical relation between area and volume and their respective linear dimensions.

Studies by Lovell, Healey, and Rowland (1962); Lovell and Ogilvie (1961); and Lunzer (1960b)--which employed items similar to those used by Piaget et al. (1960)--generally supported their conclusions regarding the development of measurement concepts. On the other hand, whereas these studies implied that conservation is a prerequisite for measurement, Bearison (1969) used measurement operations to teach children to conserve. Nonconservers were provided with experiences in which they compared two quantities of liquid in terms of the number of identical beakers containing the two quantities. Bearison (1969) concluded that:

> *The effects of training facilitated the conservation of continuous quantity and transferred to the conservation of area, mass, quantity, number, and length. The explanations offered for conservation by the trained conservers were identical to those elicited from a group of "natural" conservers, and the effects of conservation were maintained over a 7-month period (p. 653).*

In another set of studies integrating counting and conservation concepts Carpenter (1971 b) and Wohlwill and Lowe (1962) found that simply counting the elements in the conservation of discrete object problems does not substantially improve performance. However, Almy, Chittenden, & Miller (1966) and Greco, Grize, Papert, & Piaget (1960) found that "conservation of number" (invariance of the number assigned to a set of discontinuous elements under a reversible transformation) precedes the standard equivalence conservation, which involves invariance of the relation between equal quantities.

Whereas Piaget et al. (1960) considered both the empirical procedures applied directly to the domain and the application of the measurement function, two Soviet researchers, P. Ya. Gal'perin and L. S. Georgiev (1969), have concentrated their efforts on the application of the measurement function, especially the role of the unit in defining the function. They administered a series of measurement problems to a group of Soviet kindergartners in which the children were asked to measure and compare quantities of rice in various situations. Based on these studies, Gal'perin and Georgiev (1969) concluded that young children taught by traditional methods have a number of serious misconceptions regarding the measurement process because of a lack of a basic understanding of a unit of measure. They found young children to be indifferent to the size and fullness of a unit of measure and to have more faith in direct visual comparison of quantities than in measurement by a given unit.

In a replication of the Gal'perin and Georgiev investigation with American first graders, Carpenter (1971a) found responses similar to those in the Soviet study. However, based upon the results on four additional items and a different interpretation of the results of the Soviet items, the conclusion was made that young children are not indifferent to the size and fullness of units of measure; but just as in Piaget's conservation problems, they are only capable of making one-dimensional comparisons and therefore do not focus on both unit size and number of units at the same time. Furthermore, Carpenter (1971a) hypothesized that young children do not rely primarily on visual comparisons, as both Piaget and Gal'perin and Georgiev have concluded, but rather they respond on the basis of the last stimulus available, be it visual or numerical.

Carpenter's (1971a) investigation also raised questions regarding the role of equivalence and order relations in conservation and measurement problems. Measurement problems in which equal quantities were made to appear unequal by measuring them with different sized units of measure were significantly more difficult than similar items in which unequal quantities were made to appear equal. Similar results were found favoring inequality when the same unit was used to measure and compare two different quantities. These results, which suggest that with regard to certain aspects of the measurement process a stable concept of nonequivalence may precede a stable concept of equivalence, run counter to the above logical construction of these concepts in the definition of the measurement function.

In his basic works on number and measurement (Piaget 1952, Piaget et al. 1960), Piaget does not differentiate between items employing different relations between sets. He has attempted to assess the child's conception of number, length, weight, etc. of a single quantity and has used equivalent sets as an experimental convenience. Elkind (1967), Van Engen (1971), and Wohwill and Lowe (1962) have questioned Piaget's procedure of using what Elkind (1967) calls "conservation of equivalence" tasks to assess conservation of number, length, weight, etc. For example, Elkind (1967) hypothesized that "identity conservation" (invariance of a quantitative attribute --e.g., numerousness, weight, volume--under a reversible transformation) precedes equivalence conservation. This hypothesis has been supported in studies by Hooper (1969) and McMannis (1969), while a study by Northman and Gruen (1970) found no differences between the two types of conservation.

In a study directly testing the effect of equivalence and order relations on performance on conservation items, Zimiles (1966) found no significant difference in difficulty between conservation tasks using equivalent sets of discrete objects and conservation tasks using non-equivalent sets of discrete objects in which the direction of the non-equivalence appeared to be reversed after the transformation. However, there was evidence that a substantial amount of individual inconsistency of performance between items could be attributed to differences in equivalence and nonequivalence conditions.

Steffe and Johnson (1971) also found that items which contained equal numbers of items in the sets to be compared demanded different abilities than items that employed the same number of items in both sets. In several other studies both equivalence and nonequivalence items have been administered. Although these studies were not designed to test for differences between the two types of tasks, their results were examined to determine whether differences did in fact exist. Analysis of individual items in studies by Carey and Steffe (1968) and Harper and Steffe (1968) indicated no clear-cut differences in difficulty between equivalence and nonequivalence items. On the other hand, in a study of conservation of discontinuous quantity with children between 2 and 4 years old, Piaget (1968) found a significantly greater number of correct answers in non-equivalence situations. Beilin (1968) and Rothenberg (1969) also reported significantly more correct answers to problems in which the relations between sets were nonequivalence; however, their tasks were not traditional conservation problems, and experimental variables appeared to favor the nonequivalence situations.

From a slightly different perspective, the results of three studies in which the type of inference required of the children rather than the relation between the sets being compared was investigated (Beilin 1964, Carey and Steffe 1968, and Griffiths, Shantz, and Sigel 1967), indicate that problems that require judgements of equality are more difficult than problems that require judgments of inequality. In one favoring equivalence Uprichard (1970) found that treatments in which children learned to classify sets on the basis of equivalence were mastered more quickly than treatments in which children classified sets on the basis of "greater than" or "less than," and learning sequences that began with equivalence were more effective than sequences that began with either "greater than" or "less than."

Several factors may explain this rather mixed collection of results regarding the role of equivalence and order relations in conservation and measurement problems. First, certain of the nonequivalence problems may not have required true conservation judgments. For example, if unequal quantities are made to appear equal by measuring them with different size units so that they measure the same number of units, it is still possible to accurately compare the quantities on the basis of unit size with no reference to the previous state. Second, although there does not appear to be any difference in difficulty between equivalence and nonequivalence problems with discrete objects, there is some evidence that suggests that equivalence-nonequivalence differences may exist for problems comparing continuous quantities where precise judgments of equality are more difficult than judgments of inequality. Fleishmann, Gilmore, and Ginsburg (1966) and Smedslund (1966) found that a number of young children (as many as 20%) fail to maintain choices of equality even when no apparent conflict is introduced.

Purpose and Procedures

The major purposes of this study were (1) to assess the degree to which young children possess the logical structures to assimilate and apply information from measurement processes and (2) to identify some of the factors involved in the development of measurement and conservation. Specifically, the question as to whether conservation and measurement failures are primarily the result of a dependence on perceptual cues, the order of the cues or an interaction of the two was investigated. That is, an attempt was made to determine whether young children respond differently to visual and numerical cues in conservation and measurement problems or whether they simply respond to the last cue available to them.

Another purpose of the study was to determine the role of equivalence and order relations in children's performance on conservation and measurement problems. Three different combinations of order and equivalence relations were studied.

1. Equivalence: Equal quantities were transformed to appear unequal.

2. Nonequivalence I: Unequal quantities were transformed so that the dominant dimension in each quantity (height of the liquid in conservation problems--number in measurement problems) was equal.

3. Nonequivalence II: Unequal quantities were transformed so that the direction of the inequality appeared to be reversed.

For most Nonequivalence I problems the correct relation between quantities could be determined from the distracting cues by simply focusing on the appropriate dimension (for example by focusing on the size of the unit rather than the number of units). To determine whether any possible differences favoring Nonequivalence I were simply the result of this sort of pseudo conservation, differences between measurement problems in which it was not possible to visually distinguish the larger unit were assessed.

Whether recognizing that an increase in one dimension of a quantity may be compensated for by a decrease in another dimension (when holding the quantity constant) is important in young children's conservation judgments was also investigated. Piaget (1952) asserted that this recognition of compensating relationships is a significant factor in the development of conservation. By contrasting performance on measurement problems in which it was possible to visually distinguish this compensating relationship to problems in which it was not, information concerning the importance of this factor for young children's conservation judgments was obtained.

Finally, young children's understanding of the following basic measurement concepts was investigated.

1. Quantity A is equivalent to quantity B if and only if $\mu(A) = \mu(B)$, and quantity A is less than quantity B if and only if $\mu(A) < \mu(B)$.

2. In order to compare quantities on the basis of measurement, the same measurement function (the same unit) must be used to measure both quantities.

3. When equivalent quantities are measured with different units, an inverse relation exists between unit size and the number of units.

In order to conduct the investigation, the following items were administered to a group of 129 first and second graders.

1. Conservation of continuous quantity.

   Equivalence. The child was shown two identical glasses containing equal amounts of water and was asked to compare the amount of water in the two glasses. If he said that there was more water in one of the glasses, some water was poured from this glass into the other glass; and this process was repeated until the child agreed that there was the same amount of water in the two glasses. Then one of the glasses of water was poured into a taller, narrower glass, and the child was again asked to compare the amounts of water.

   Nonequivalence I. The child was shown two identical glasses containing unequal amounts of water and was asked to compare the amounts of water in the two glasses. Then the glass containing the smaller amount of water was poured into a taller, narrower glass such that the height of the water was the same as the height of water in the glass containing more water, and the child was again asked to compare the two amounts of water.

   Nonequivalence II. The child was shown two identical glasses containing unequal amounts of water and was asked to compare the amount of water in the two glasses. Then the glass containing the smaller amount of water was poured into a taller, narrower glass such that the height of the water was higher than the height in the glass containing more water, and the child was again asked to compare the two amounts of water.

2. Measurement with visibly different units.

   Equivalence. The child was shown two glasses containing equal amounts of water and was aked to compare the amounts of water in the two glasses. If he said that there was more water in one of the glasses, some water was poured from this glass into the other glass; and this process was repeated until the child agreed that there was the same amount of water in the two glasses. Then the water in each glass was measured into two opaque containers using visibly different units of measure so that one glass of water measured three units and the other measured five. Then the child was again asked to compare the two amounts of water.

Nonequivalence I. The child was shown two glasses containing unequal amounts of water and was asked to compare the amount of water in the two glasses. Then the water in each glass was measured into two opaque containers using visibly different units of measure such that both glasses measured three units. Then the child was again asked to compare the two amounts of water.

Nonequivalence II. The child was shown two glasses containing unequal amounts of water and was asked to compare the amount of water in the two glasses. Then the water in each glass was measured into two opaque containers using visibly different units of measure so that the greater quantity of water measured three units and the other measured four. Then the child was again asked to compare the two amounts of water.

3. Measurement with indistinguishable different units.

Equivalence. This task was the same task as the equivalence task in measurement with visibly different units, except the smaller unit appeared larger. One glass measured five units and the other measured four.

Nonequivalence I. This task was the same task as the non-equivalence I task in measurement with visibly different units, except the smaller unit appeared larger. Both glasses measured four units.

Nonequivalence II. This task was the same task as the non-equivalence II task in measurement with visibly different units, except the smaller unit appeared larger. The greater quantity of water measured six units and the other measured seven.

4. Measurement of unequal-appearing quantities with the same unit.

Equivalence. The child was asked to compare two equal quantities of water in two different-shaped containers, one tall and narrow and the other short and wide (i.e., the final state in the equivalence task of conservation of continuous quantity). Then the water in each glass was measured into two opaque containers using the same unit (each glass measured four units) and the child was asked to compare the two amounts of water.

Nonequivalence II. The child was asked to compare two unequal amounts of water in the two different-shaped containers (the final state in the nonequivalence II task of continuous quantity). Then the water in each container was measured into two opaque containers using the same unit (the glass that appeared to have more water measured four units and the other measured five), and the child was again asked to compare the two amounts of water.

5. Measurement with the same unit into apparent inequality.*

Equivalence. Using the same unit of measure, four units of water were measured into two different-shaped containers and the child was asked to compare the two quantities of water.

Nonequivalence II. Using the same unit of measure, five units of water were measured into a short, wide container and four units were measured into a tall, narrow container and the child was asked to compare the two quantities of water.

In order to keep the number of tasks administered to each child reasonable, items were split into two groups and each group was administered to a different set of children. Sixty-one children in Part A received all three conservation problems and both sets of problems in which quantities were measured with two different units. All three sets of problems were administered with each of the three relations. Sixty-eight children in Part B received all the measurement problems with Equivalence and Nonequivalence II relations.

Thus, the problems in Part A fit a 3 x 3 repeated measures design where the factors were Problem Type (continuous quantity conservation and the two measurement problems with different units) and Relations (Equivalence, Nonequivalence I, and Nonequivalence II). The problems in Part B fit a 2 x 4 repeated measures design where the factors were Problem Type (all four measurement problems) and Relations (Equivalence and Nonequivqlence II).

The hypotheses of interest tested in Part A are as follows.

H_1: There is no significant difference between performance on Equivalence and Nonequivalence II items for any of the problem types.

H_2: There is no significant difference between performance on Equivalence and Nonequivalence I measurement problems in which the larger unit of measure is not visually distinguishable.

H_3: There is no significant difference between performance on Equivalence and Nonequivalence I items for conservation problems or for measurement problems in which the larger unit is visually distinguishable.

H_4: There is no significant difference between performance on conservation problems and corresponding measurement problems.

H_5: There is no significant difference between performance on measurement problems in which it is possible to visually distinguish the larger unit and those in which it is not.

H_6: Neither mean performance nor any of the above contrasts are significantly affected by grade, sex, or the order in which the items were administered.

*These tasks are simply the tasks in (4) with the stimuli appearing in a different order.

Hypotheses H1, H5, and H6 were also tested in Part B. In addition, the following hypotheses were added:

H_7: There is no significant difference in performance between measurement problems in which correct visual cues are followed by distracting numerical cues and corresponding problems in which correct number cues are followed by distracting visual cues.

H_8: There is no significant difference in performance between measurement problems in which the correct measurement cues appear before distracting visual cues and those in which they appear after the distracting visual cues.

Subjects. This study was run over a nine-day period in the spring of 1971 in a predominantly rural community near Madison, Wisconsin, with a population of about 4,000. The subjects (*Ss*) for the study were selected from three of the five first grade classes and two of the five second grade classes in one of the two elementary schools serving the community. The sample, which included all students in the five classes except three who were absent on the testing days, consisted of 75 first graders and 54 second graders. The age range of the first graders was 6 years, 5 months to 9 years, 8 months with mean age 7 years, 5 months; and the range of the second graders was 7 years, 7 months to 9 years, 5 months with mean age 8 years, 4 months.

Procedures. *Ss* were randomly assigned to two groups, 61 *Ss* to Part A and 68 to Part B. Each *S* within each group received the same basic set of problems; however, the order of the problems was randomized for each *S*.

All items were administered in a small room apart from the classroom by one experimenter (E), a stranger to the *Ss*. The *S* sat at a table opposite the E. Procedures and protocols were kept as consistent as possible between items; however, certain procedures were randomly varied between *Ss* in order to control for responses based on experimental variables.

Piaget (1968) and Siegel and Goldstein (1969) found that young children tend to respond to the last choice available to them in a conservation problem. Thus, if the E said, "Is there the same amount of water in the two cups or does one have more?" the *S* may respond that one has more because "more" was the last choice given to him. Therefore, some of the *Ss* were asked, "Is there the same amount of water in the two cups or does one have more?" and the others asked, "Does one cup have more water in it or is there the same amount in each cup?". For each *S* the "same-more" order was the same for all problems.

For some *Ss* the smaller quantity was always measured first in nonequivalence problems, and for others the larger quantity was always measured first. Both of these variations were randomly assigned to *Ss*.

Problems were administered in two sittings. *Ss* in Part A received five problems the first day and four several days later, and *Ss* in Part B had four problems the first day and four the second. Although reasons for responses were solicited and recorded, answers were judged correct or incorrect without regard to the explanations given.

Analysis. Item totals, reasons for responses, and types of errors were recorded for each item. The following categories were used to classify reasons for correct responses:

1. Reversibility: If the quantities were transformed back to their former state, they would again appear in the correct relation (equal or unequal).

2. Statement of operation performed: The water was just poured into a different container and this did not change the relation between the quantities.

3. Addition--subtraction: Nothing was added or taken away.

4. Compensation, proportionality: The liquid was higher but the container was narrower. One measured more units but the units were smaller.

5. Sameness of quantity: It's the same water.

6. Reference to the previous state: They were the same before when the water was in identical glasses.

7. No reason, unclassifiable: No reason was given or an incomprehensible reason was given.

Incorrect responses were sorted into two broad categories.

1. Dominant dimension: *S*s incorrectly chose (a) the taller container of water or (b) the quantity that measured the greater number of units.

2. Secondary dimension: *S*s incorrectly chose (a) the wider container or (b) the quantity measured with the larger unit.

Hypotheses were tested using a multivariate analysis of variance program of J.D. Finn (1967). In this program, analysis is conducted using single degrees of freedom-planned contrasts, and freedom is allowed to specify the contrasts of interest. This flexibility especially suited the purpose of this study in which specific contrasts, rather than overall differences between factors, were of interest.

The Finn program yields standard errors for each of the variables and also estimates the magnitude of the effects for the specific contrasts and their standard errors. Thus, 95% confidence intervals have been plotted for each of the problems and for each of the significant contrasts.

Based on the results of previous research it was predicted that there would be no significant difference between performance on Equivalence and corresponding Nonequivalence II problems. However, Nonequivalence I problems were expected to be easier than problems employing the other two relations except in the set of measurement problems in which the larger unit was not visually distinguishable. Since the Nonequivalence I relation was expected to operate differently than the other two relations, contrasts between problems involving the Nonequivalence I relation were conducted independently of contrasts between problems involving the other two relations.

In both Part A and Part B, four MANOVAs were employed. The first contained all contrasts that were predicted to have no significant effect, Equivalence—Nonequivalence II contrasts, and all contrasts between measurement problems in which the larger unit was not visually distinguishable. The second MANOVA contained all contrasts that were expected to be significant and all contrasts for which there was insufficient prior evidence to make a prediction. Thus, the second MANOVA contained contrasts between problem types and contrasts between Nonequivalence I and the other two relations. The third and fourth MANOVAs tested for effects due to grade, sex, and order in which the items were administered. What this partitioning effectively did was to hypothesize a model, to test the goodness of fit of this model (the first MANOVA), to test whether the parameters of the model are nonzero (the second MANOVA), and to test whether the model or the parameters of the model are significantly influenced by grade, sex, or the order of administration of the items (the third and fourth MANOVAs).

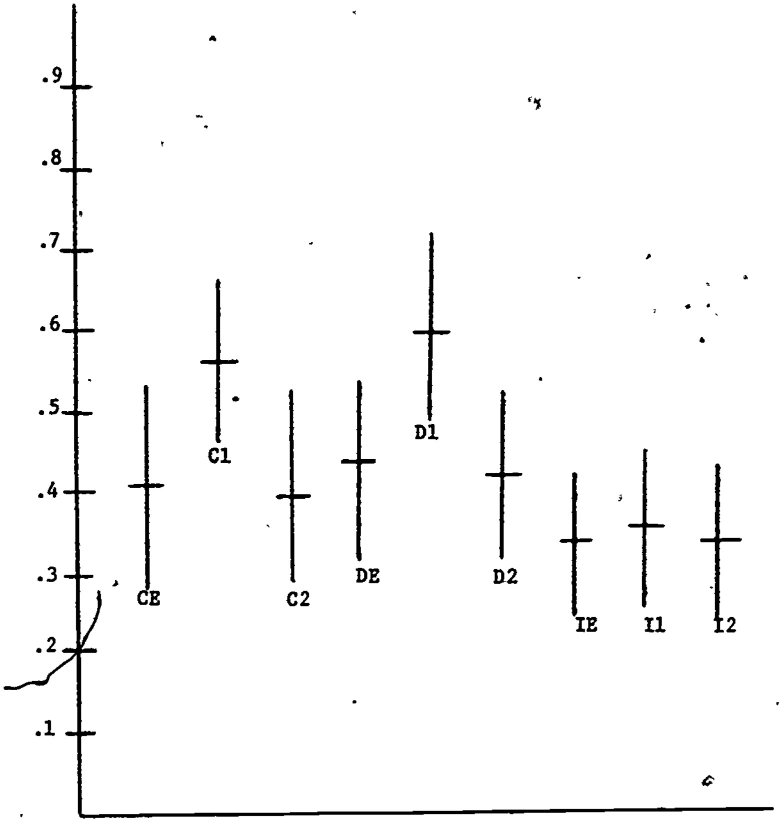
Testing for the effect of order presented certain problems. Since the order of items was randomized, every *S* received a different order of items, eliminating the feasibility of partitioning into each distinct order. Therefore, the effect of the order of the items was determined by a procedure proposed by Zimiles (1966). He found that the first item administered often significantly influenced performance on all subsequent items. *S*s administered easier first items performed better on all subsequent items than *S*s administered a more difficult first item. No differences were found, however, due to variations in the second item administered. Thus, *S*s were partitioned into order groups based on the first item they received.

Results

Part A. The results of individual items in Part A, the reasons given for responses, and the types of errors are summarized in Table 1. The means for individual items surrounded by 95% confidence intervals have been plotted in Figure 1. Since the individual items are scored on a 0-1 basis, the mean can be interpreted as representing the fraction of *S*s correctly responding to the item. Similarly the confidence intervals can be interpreted in terms of percents. For example, there is a 95% probability that between 27% and 56% of the population would respond correctly to the Equivalence conservation item.

There was very little diversity in the reasons given for correct responses. Practically all the *S*s either referred to the previous state of the quantities or noted the compensating relationship between unit size and the number of units or between height and width. Comparisons between reasons given by *S*s who were successful on the items in which it was possible to distinguish the compensating relationship between unit size and number of units but were unsuccessful on problems in which it was not are enlightening. Seven of the eight *S*s who either (1) correctly answered at least two of the measurement problems in which the larger unit was distinguishable but none of the problems in which it was not or (2) correctly answered all three of the problems in which the larger unit was distinguishable and at most one of the problems in which it was not, gave compensation as the reason for at least one of their correct responses.

Figure 1. 95% Confidence Intervals for Items in Part A



C = Conservation of continuous quantity
D = Measurement with visibly different units
I = Measurement with indistinguishably different units
E = Equivalence
1 = Nonequivalence I
2 = Nonequivalence II

The tests of the hypotheses are summarized in Table 2 and Table 3.

Table 2
MANOVA--Relation Contrasts for Part A

Source	df	MS	F	P
Multivariate	4,24		.0769	.9887
H1a: CE = C2	1	.0164	.0830	.7755
H1b: DE = D2	1	.0164	.1021	.7518
H1c: IE = I2	1	.0164	.0450	.8336
H2: IE = I1	1	.0164	.0984	.7563

Degrees of freedom for error = 27
C = Conservation
D = Measurement with visibly different units
I = Measurement with indistinguishably different units
E = Equivalence
1 = Nonequivalence I
2 = Nonequivalence II

Table 3
MANOVA--Problem-Type and Equivalence--Nonequivalence I Contrasts

Source	df	MS	F	P
Multivariate	5,23		30.5264	.0001
I = 0	1	6.7776	67.1332	.0001
H5: D = I	1	.4376	3.6255	.0677
H4a: DE + D2 = CE + C2	1	.0164	.1180	.7339
H4b: D1 = C1	1	.0164	.0608	.8072
H3: DE + CE = D1 + D2	1	1.8074	11.6189	.0021

Degrees of freedom for error = 27
C = Conservation
D = Measurement with visibly different units
I = Measurement with indistinguishably different units
E = Equivalence
1 = Nonequivalence I
2 = Nonequivalence II

These results indicate that there are no significant differences between Equivalence and Nonequivalence II relations for any of the problems tested. There is a significant difference between Nonequivalence I and the other two relations except in the case of the measurement problems in which the larger unit is not visually identifiable. No significant differences were found between the conservation and measurement problems or between the two types of measurement problems.

These results are summarized in the model in Table 4. The parameters of the model are all positive and significant except for θ. θ was included in the model because (1) it approaches significance and (2) this effect was significant in Part B (see below). Ninety-five percent confidence intervals for each of the parameters have been plotted in Figure 2.

Figure 2. Confidence Intervals for Parameters of the Model in Part A

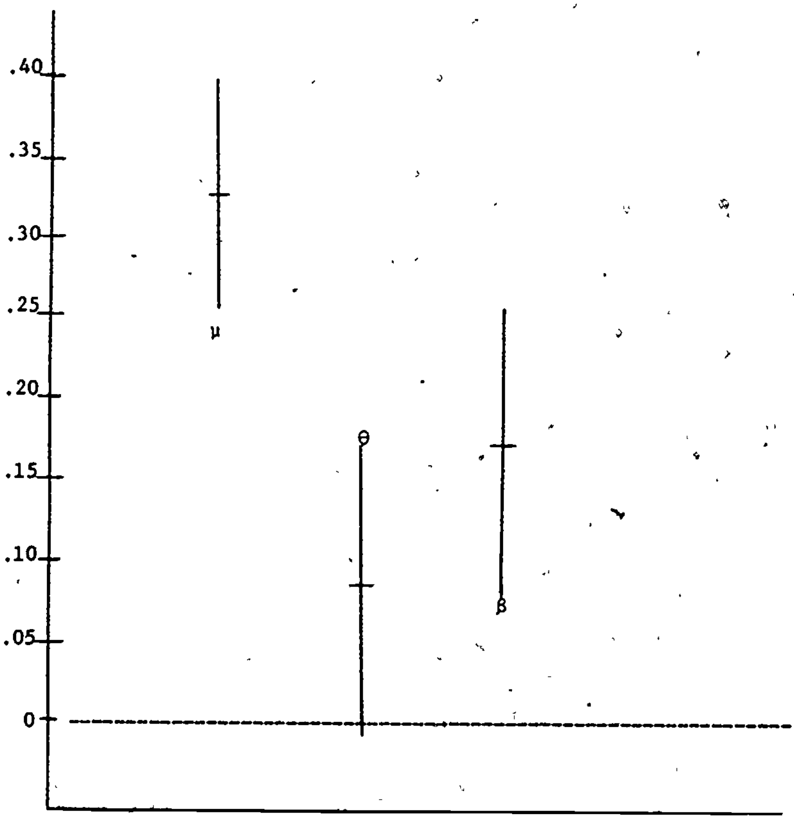


Table 4
Item Means for Problems In Part A

Problem type	Relation: Equivalence	Nonequiv. I	Nonequiv. II
Conservation continuous quantity	$\mu + \theta$*	$\mu + \theta$* $+ \beta$	$\mu + \theta$*
Measurement with distinguishably different units	$\mu + \theta$	$\mu + \theta$* $+ \beta$	$\mu + \theta$*
Measurement with indistinguishably different units	μ	μ	μ

*θ is not significant in Part A.

Significant differences between grades were found for overall means ($p = .0001$), but no significant effect due to grade level was found for any of the hypotheses tested ($p > .42$). No significant differences were found due to sex or order in which the items were administered ($p > .16$).

Part B. The results of individual items in Part B, the reasons given for responses, and the types of errors are summarized in Table 5.

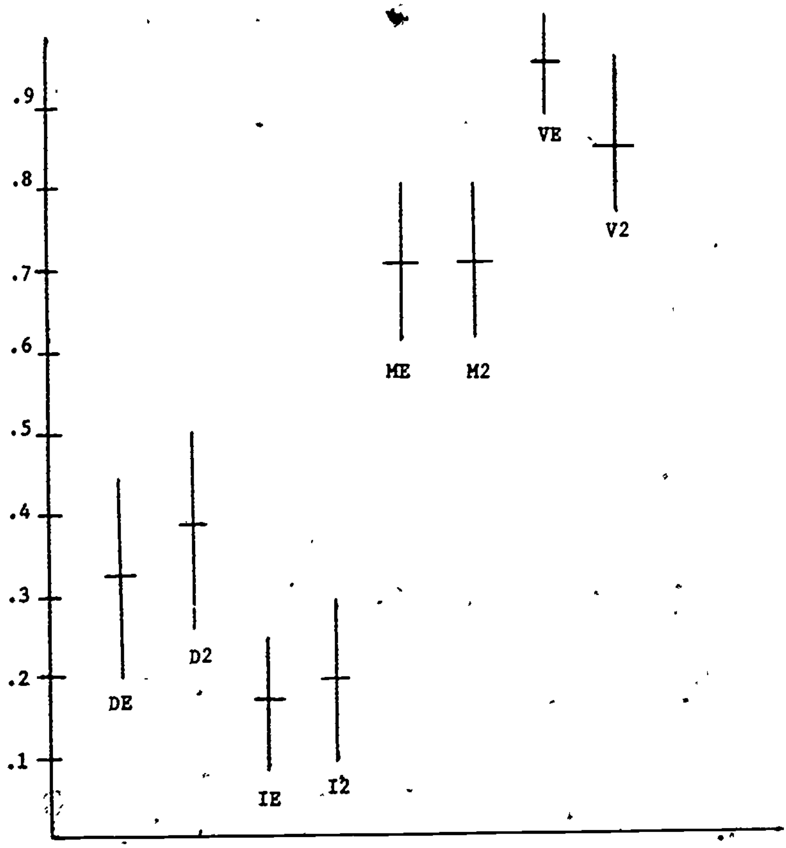
Table 5
Number of Subjects in the Major Response Categories in Part B

Item	DE	D2	IE	I2	ME	M2	VE	V2
Total Correct	22	26	11	13	47	48	64	58
Reason for correct response								
Reversibility	0	0	0	0	0	0	0	0
Statement of operation performed	0	0	0	0	0	0	0	0
Addition-Subtraction	0	0	0	0	0	0	0	0
Compensation, proportionality	5	10	0	0	0	0	0	0
Sameness of quantity	0	0	1	0	0	0	0	0
Reference to previous state	14	11	6	8	45	46	59	55
No reason given or unclassifiable reason given	3	5	4	5	2	2	5	3
Total incorrect	46	42	57	55	21	20	4	10
Type of error								
Taller container or greater number of units	46	40	57	53	21	17	4	7
Wider container or larger unit	0	2	0	2	0	3	0	3

D = Measurement with visibly different units
I = Measurement with indistinguishably different units
M = Measurement with the same unit into apparent inequality
V = Measurement of unequal-appearing quantities into the same unit
E = Equivalence
2 = Nonequivalence II

and the means for individual items surrounded by 95% confidence intervals have been plotted in Figure 3.

Figure 3. Confidence Intervals for Items in Part B



D = Measurement with visibly different units
I = Measurement with indistinguishably different units
M = Measurement with the same unit into apparent inequality
V = Measurement of unequal-appearing quantities with the same unit
E = Equivalence
2 = Nonequivalence II

As in Part A, virtually all of the reasons for correct responses fall into two categories; and except for the problems in which quantities were measured with distinguishably different units, virtually all the correct responses were based on reference to the previous state. Two of the five *Ss* who correctly answered both of the problems in which the larger unit was distinguishable but neither of the problems in which it was not gave compensation as the reason for their response.

Only one *S* missed every item. Another *S* completely ignored the number cues, even though he successfully counted the number of units; consequently, he missed all the problems in which quantities were measured with the same unit but answered correctly the items in which quantities were measured with different units. A third *S* who was in the "more-same" protocol group responded "same" to every item. On the measurement problem with indistinguishably different units, only two of the *Ss* were able to use the information from the measurement operation to correctly identify the larger unit. The rest were unable to apply the inverse relationship between unit size and number of units to this problem and simply responded incorrectly on the basis of the unit that looked larger. Between 85% and 89% of the *Ss* gave the same response to corresponding Equivalence and Nonequivalence II problems.

The analysis summarized in Table 6 indicates that there is no significant difference between Equivalence and Nonequivalence II relations.

Table 6
MANOVA--Relation Contrasts for Part B

Source	df	MS	F	P
Multivariate	3,37		1.6369	.1970
IE = I2	1	.0000	.0000	1.0000
DE = D2	1	.2353	1.9335	.1721
ME = M2	1	.0000	.0000	1.0000
VE = V2	1	.5294	4.6109	.0379

Degrees of freedom for error = 40
D = Measurement with visibly different units
I = Measurement with indistinguishably different units
M = Measurement with the same unit into apparent inequality
V = Measurement of unequal-appearing quantities with the same unit
E = Equivalence
2 = Nonequivalence II

Consideration of the univariate analysis indicates that the contrast between problems in which unequal-appearing quantities are measured with the same unit approaches significance at the .01 level adopted in this study and would be significant if a .05 level had been adopted. None of the other contrasts even approach significance. The analysis in Table 7 indicates that there is a significant difference between each of the four types of measurement problems in Part B.

These results are summarized in the model in Table 8. The parameters of the model are all positive and significant. Ninety-five per cent confidence intervals for each of the parameters have been plotted in Figure 4.

Figure 4. Confidence Intervals for Parameters in Part B

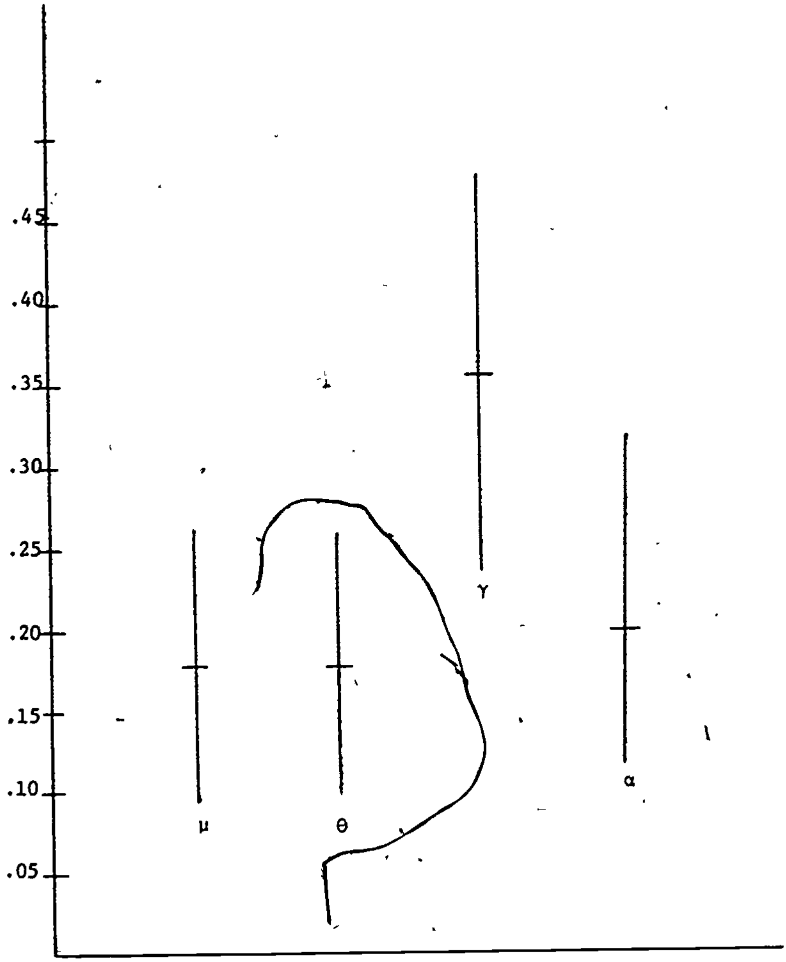


Table 7
MANOVA--Problem-Type Contrasts for Part B

Source	df	MS	F	p*
Multivariate	3,38		70.8403	.0001
H5: D = E	1	2.1176	20.3702	.0001
H6: D = M	1	8.4706	33.6859	.0001
H7: M = V	1	2.4853	19.5404	.0001

Degrees of freedom for error = 40
D = Measurement with visibly different units
I = Measurement with indistinguishably different units
M = Measurement with the same unit into apparent inequality
V = Measurement of unequal-appearing quantities with the same unit

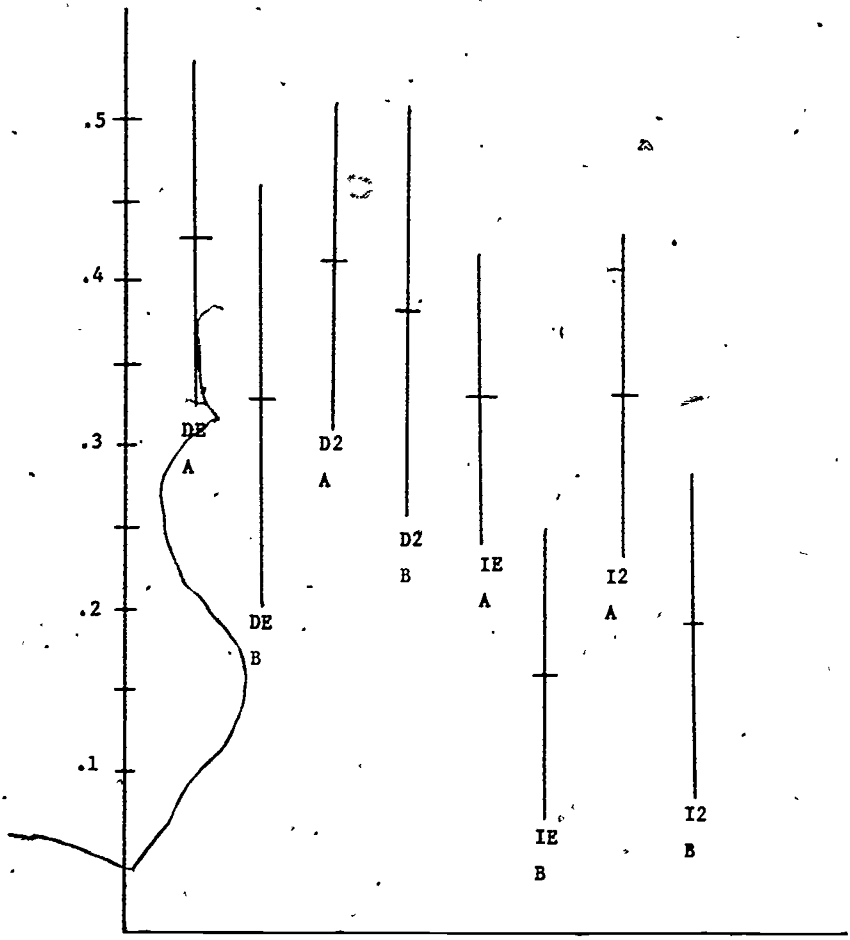
Table 8
Means for Items in Part B

Problem Type	Relation: Equivalence	Relation: Nonequiv. II
Measurement with distinguishably different units	$\mu + \theta$	$\mu + \theta$
Measurement with indistinguishably different units	μ	μ
Measurement into different-shaped containers	$\mu + \theta + \gamma$	$\mu + \theta + \gamma$
Measurement from different-shaped containers	$\mu + \theta + \gamma + \alpha$	$\mu + \theta + \gamma + \alpha$

No significant differences were found due to grade, sex, or the order in which the items were administered. Sex and order effects approached significance ($p = .02$) for the contrast between the two measurement problems in which a single unit was used, but did not even approach significance for any of the other contrasts ($p > .13$).

A and B comparisons. Four items were given in both parts of the study. The 95% confidence intervals for corresponding items do intersect (Figure 5); however, comparison of corresponding item means indicates that the fact that differences between measurement problems using visibly different units and those using indistinguishably different units are significant in Part B but fail to reach significance in Part A can be attributed entirely to between-study differences in performance on the problems employing indistinguishably different units.

Figure 5. Confidence Intervals for Items Appearing in Both Parts A and B



D = Measurement with visibly different units
I = Measurement with indistinguishably different units
E = Equivalence
2 = Nonequivalence II

Analysis of variance for contrasts between conservation of continuous quantity and measurement problems in which quantities were measured with the same unit into apparent inequality (i.e., the final state of the conservation problems) is summarized in Table 9, indicating significant differences favoring the measurement problems. These results should be interpreted somewhat cautiously in that the two types of problems were administered in different sets of problems in the series.

Table 9
ANOVA--Conservation of Continuous-Quantity Measurement into Apparent Inequality Contrast

Source	df	MS	F
Between	1	11	14.6**
Within Cells	127	.76	

**p < .01

The results in Figure 5, however, indicate that for the four problems that were administered in both parts, performance was generally higher in the part containing the conservation problems; so the danger of interaction with other problems favoring the measurement problems is probably not too great.

Furthermore, in Part A no significant difference was found between conservation problems and corresponding problems in which quantities are measured with different units. In Part B problems in which quantities are measured with the same unit were found to be significantly easier than corresponding problems in which two units are employed. The combination of these results confirms that problems in which quantities are measured with the same unit are easier than corresponding conservation problems.

Summary and Conclusions

It appears that it is not simply the perceptual properties of the stimuli that produce errors in conservation problems. There is no significant difference in difficulty between conservation problems and corresponding measurement problems in which the distracting cues are numerical. The position of Piaget (1952, 1960), Bruner, Olver, and Greenfield (1966) and others that young children are highly dependent on perceptual properties of events and that conservation problems occur because the immediate perceptual properties of the conservation problems override the logical properties that imply conservation, has been based on tasks in which distracting visual cues always appeared last. The results of the current investigation, however, demonstrate that misleading numerical cues produce the same errors as misleading visual cues.

The failure of young children to respond primarily on the basis of visual cues is even more striking in the contrast between conservation problems and the problems in which quantities are measured into apparent inequality and the contrast between the problems in which quantities are measured with distinguishably different units and the problems in which quantities are measured into apparent inequality. The problems measuring quantities into apparent inequality, in which correct measurement cues are followed by misleading perceptual cues, are significantly easier than either corresponding conservation problems, in which both sets of cues are visual, or corresponding problems in which quantities are measured with different units, where correct visual cues are followed by incorrect numerical cues.

These results, which could be interpreted to imply that numerical modes dominate visual modes, should be regarded with some caution. Zimiles (1963) has suggested that conservation failures may result from *S*s basing their judgments on the E's manipulations of the quantities being compared. For example, if two rows of blocks which the *S* has judged equivalent when they are arranged in one-to-one correspondence are spread out, the *S* says that the longer row has more because the act of spreading the blocks out implies to him that the length of the rows is the dimension he is being asked to compare.

In the current investigation, the experimental procedures emphasize the measurement cues, which means that the correct choice is emphasized in the problems employing a single unit of measure but the incorrect choice is emphasized in the problems employing different units of measure and the conservation problems.

Thus, it appears that the most significant factor in determining which cues young children attend to is the order in which the cues appear. Problems in which correct cues appear last are significantly easier than corresponding problems in which correct cues are followed by misleading cues. As noted above, however, the order of the cues was not the only factor that was found to affect responses.

In general, there does not appear to be any significant difference between conservation and measurement problems employing Equivalence relations and corresponding problems employing Nonequivalence II relations. Nonequivalence I problems are significantly easier than corresponding problems employing Equivalence or Nonequivalence II relations except in problems in which it is not possible to identify the larger unit. These results imply that the relation between quantities being compared does not affect performance, and the Nonequivalence I problems are easier simply because they do not require genuine conservation, since accurate comparisons can be made from the final states of the quantities.

Measurement operations have some meaning for the majority of students in the first and second grades. By the end of the first grade, virtually all students recognize that quantities are equal if they measure the same number of units and quantity A is greater than quantity B if A measures more units than B. Only 3 of the 129 *S*s tested did not respond to any questions on the basis of measurement cues; and only 2 of the 3 definitely ignored the measurement cues. The other *S* simply responded "same" to all problems.

This does not mean, however, that first- and second-grade students have totally correct measurement concepts or are able to accurately apply measurement processes. As few as 25% of the *S*s tested completely understood the importance of using a single measurement function, and only 6% were able to discover the relation between measurement functions from the results of the measurement operations. Only 70% of the *S*s were able to use measurement results if they were followed by conflicting visual cues. Only 59% of the *S*s demonstrated any knowledge that variations in unit size affected measurement results, and as few as 40% of the *S*s were able to apply this knowledge to problems in which quantities were measured with different units. This figure dropped to 25% when the larger unit was not visually distinguishable, and only 6% of the *S*s were able to use the results of measurement operations to determine the larger unit when it was not visually apparent.

The conclusion that by the end of first grade virtually all children, even those in Stages I and IIA, have some concept of measurement appears to contradict Piaget's (1960) conclusions that measurement concepts do not begin to appear until Stage IIB. This apparent conflict is due to the fact that Piaget employed less structured measurement tasks. In order to have any measurement cues to respond to, *S*s had to measure themselves. In the current investigation the measurement cues were forced upon the *S*s; therefore, even *S*s in the earliest stages had number cues to guide or distract their reponses.

The results of the two sections of this study with respect to the importance of recognizing the compensating relation between dimensions are ambiguous. Significant differences between the problems in which it is possible to distinguish the larger unit and those in which it is not were found in Part B but not in Part A. In Part A, however, about 7% of the *S*s tested did find the problems in which the larger unit was distinguishable easier than the problems in which it was not, and the pooled results indicate there exist significant differences. Consideration of this fact and examination of the confidence intervals for the parameter Θ indicate that probably at least 10% of the population sampled require that the distracting cues contain compensating relations in order to conserve. These conclusions should be regarded with some caution, however, since the discrepancy between the results in Parts A and B indicates that there may be some interaction between tasks administered to the same *S* that affects the parameter Θ.

In general there were no significant differences due to sex, order of items, or protocol variations. The fact that only one *S* in the entire investigation consistently responded either "more" or "less" to all problems indicates that by the end of the first grade few children still respond to conservation problems on the basis of the last alternative offered to them.

Thus, of the factors under consideration in this study it appears that centering on a single dominant dimension is the major reason for most conservation and measurement failures and the development of conservation and measurement concepts can be described in terms of increasing ability to decenter.

In the earliest stage children respond on the basis of a single immediate, dominant dimension. The dimension may be either visual or numerical, depending on the problem. In the next stage children are capable of changing dimensions, but in each problem they still focus on a single dimension. In this stage some children are capable of pseudo conservation and correctly solve Nonequivalence I tasks but not Equivalence or Nonequivalence II tasks. Children in this stage occasionally conserve by focusing on the earlier state but are incapable of simultaneously considering the immediate state of the quantity and the state prior to the transformation and deciding which set of cues provide a legitimate basis for comparison. They generally explain their responses by referring to the prior state and seem to purposely ignore the current state of the quantity. Children around this stage probably show the most gains in conservation training research, but the gains may be in a very narrow sense. It seems likely that most training simply serves to redirect the object of the child's centering without providing him with the flexibility of thought that is necessary for progression to later stages. Thus the child may learn to conserve by simply centering on a different aspect of the problem but still lack the flexibility of thought that conservation implies.

Finally, children gain the flexibility to consider several conditions of a quantity simultaneously and can choose the condition that provides a rational basis for comparison; however, it is not until a later stage that they are able to consider the consequences of the comparisons between different states and use the information from both conditions to discover the correct relation between the sizes of different units based on the number of units two comparable quantities measure.

KENNETH LOVELL

Summary and Implications

The research papers that have been presented represent a concerted effort to investigate the thinking of children in the age range 4 to 8 years. Five of the six papers have dealt with aspects of the training and acquisition of logical structures and the sixth with the acquisition of structures involved in conservation and measurement. Accordingly, we must begin by briefly commenting on what is known of the effects of training. Beilin (1971) has given an extensive review of the broad position on the effects of training as it was at the end of 1970, and the position at the time of writing this chapter has not materially changed. If the training of conservation or of logical operations (classification and relational skills) is considered, training does often appear to effect an improvement in performance. At least, the training appeared to work in the hands of its proponents. The Geneva workers would probably concur, although they would stress that training has no effect if some vestige of operativity is not already present (Inhelder and Sinclair, 1969). For both those who tend to give credence to the Piagetian conceptual framework and for those who do not, the problem is (as Beilin points out) to define what is meant by true operativity. If strong criteria are insisted on it is difficult to refute the Piagetian position. But if weaker criteria are used it is easier to disconfirm it. However, it would not be profitable to pursue this particular point here, vital as it may be. Rather the results of the experiments in terms of the effect of training on performance on the posttests are considered.

Even if the issue of operativity is not pursued, other issues that could influence interpretations of the experiments remain. First, it is not possible for the reader to know much about the quality of the training in a particular study. While it is true the standard procedures can be laid down, one cannot be sure (unless present) of the general atmosphere and quality of the interaction between experimenter and pupils. Second, there is the question of the size of the groups involved in the training programmes. The writer is well aware of the difficulties in selecting pupils and of the work involved. But it must be pointed out that it is possible for fluctuation to be present in data obtained from small groups. My experience leads me to suggest that when small groups are themselves randomly selected from small numbers of children, it is possible for the small groups to be different with respect to some relevant variables. In such cases, even when results are obtained at some acceptable level of statistical significance, they cannot necessarily be trusted too far. In the five papers involving training, summary information is contained in Table 1. The results of these studies, then, can be generalized only to

Table 1
Number of Children in the Samples of the Training Studies

Paper	Numbers of Children	Children Were:
Steffe and Carey	20 four-year-olds 34 five-year-olds	
Owens	23 Kindergarteners 24 First Graders	Randomly assigned to full and partial treatments
Johnson, M. L.	39 First Graders 42 Second Graders	Selected from two schools
Lesh	20, selected from a larger group used in pilot study	Randomly assigned to experimental and control groups
Johnson, D. C.	115 children of which 35 were alternates	Randomly selected from two age levels and two intelligence levels and then randomly assigned to experimental and control groups with ten per cell.

similar samples. While it is possible that the groups were different with respect to some relevant variables, there is no way that this can be ascertained. A third issue is one that was raised by Smedslund (1964). He argues that in every concrete reasoning task there should be clear distinction made between *percept, goal object and inference pattern.* The first resides in the stimulus situation as apprehended by the subject; the goal object is what the child is told to attain, such as quantity or length; while inference pattern is formed by the set of premises and conclusion, e.g., transitivity or conservation. He argues that each factor can only be studied with the other two held constant -- this being a necessary condition for the discovery of exact relations. For his study, 160 children were involved ranging in age from 4-3 to 11-4 and evenly distributed over age and sex. The tasks[1] were administered individually. The results strongly point to the fact that when the generality of concrete reasoning is studied over variations in percepts, goal objects, and inference patterns, concrete reasoning has a very limited generality over the period of

[1]These included class inclusion, multiplication of classes, reversal of spatial order, conservation of discontinuous quantity, multiplications of relations, transitivity of length, conservation of length, addition and subtractions of one unit, transitivity of discontinuous quantity.

its acquisition. Indeed, in such circumstances, it appeared to be acquired in one restricted situation at a time, and only the grosser regularities — such as differences in average item difficulty -- were observed. However, when the interactions between different inference patterns were studied, with goal objects and percepts held constant, the position was different. Such a study is, of course, of great importance, since concrete reasoning is assumed to be reflected in certain types of inference patterns. In Smedslund's investigation, the comparison between conservation and transitivity of length approached methodological perfection in the sense that only inference patterns varied, and goal objects and percepts were virtually constant. In this situation, only one child passed the test of transitivity and failed the test of conservation. Smedslund assumed that this was a case of diagnostic error and that conservation of length preceded transitivity of length.

Another issue (although not unrelated to the third) involves the developmental links between partial structures. Inhelder (1972) quotes earlier work of the Genevans into the question of whether elementary measurement of length can be helped by the application of numerical operations. Their conclusion was that interactions between numerical and ordinal ways of dealing with the problem of either judging or constructing lengths tend to produce conflict, and it was this which led to the final resolution of the measurement task, since conflict was overcome only by the child's own efforts to find "compensatory and coordinating actions." They argued that psychologically speaking, conflicts give rise to the recombination of existing partial structures in order to reestablish the equilibrium which has been destroyed, and hence, conflicts give rise to new constructions. Inhelder argues from biology, where new combinations can only take place inside what are called *reaction norms*, that new combinations in cognitive development can only occur within narrow zones of assimilating capacities. The structural levels of thought, while being at the very origin of the generation of new combinations, at the same time impose limits on the new constructions that can be produced.

Closely linked with the contents of the last paragraph is the fact that in scholastically backward, and particularly in school-educable retarded children, the attainments of concrete operational thought are extremely erratic from situation to situation (Lovell, 1966b). Both in everyday life and school situations, it would appear that their experiences bring far fewer conflicts, and/or they have very narrow zones of assimilating capacities. It is thus somewhat trite to say that given training experiences may not be assimilable to children of limited ability. There may well be, of course, something of a "chicken-egg" problem here. However, Owens quotes the study of Skypek who found the developmental pattern of cardinal number conservation was likewise erratic among pupils from lower socio economic- status homes.

These, then, are some of the issues which must be borne in mind when considering the results of the experimental studies and their educational implications.

The Research Results and the Educational Implications

The paper by Steffe and Carey is discussed first. It should be remembered that the measured intelligence of the children was normally distributed in both age groups (although both means were well above the average of 100), and the distribution of social classes to which the children belonged was also a normal one. Even on the first application of the Length Comparison Test, many children in both age groups could establish relations between two curves, especially in the case of "longer than." The remaining results are summarized in Table 2.

Table 2
Summary of Results of the Study by Steffe and Carey

Training	Results on Posttests
(Involving the interplay of language and action.)	
Instructional Sequence I.	A significant improvement in the number of correct responses, in comparing the lengths of two curves, especially in the case of the relations "shorter than" and "same length as."
Instructional Sequences II and III.	1. Little improvement in the number of correct responses in the comparison of lengths of two curves, since children came to the first application of the test fairly well able to discriminate among the relations. 2. A significant improvement in performance in respect of the Length Conservation Test: Level I. The percent of children categorized at Level I increased from 12 to 57. 3. The percent of children categorized at both Levels I and II of the Length Conservation Test increased from 8 to 37 — a statistically significant increase. 4. A significant improvement in the ability of the age groups to use both reflexive and nonreflexive properties. In the posttest some 41 percent were able to use the reflexive property and 30 percent both properties. 5. No four-year-olds were able to use the transitive property either before or after the instructional sequences, but the percent of five-year-olds able to use the transitive property increased from 16 to 31.

The variables IQ, Verbal Maturity, Age, and Social Class had no significant effect on the posttest scores in the case of the Length Comparison Test, the Length Comparison Application Test, Length Conservation Test. Levels I and II, and the test of Reflexive and Non-Reflexive properties. This argues a case that the appropriate instructional activities may profitably be undertaken with similar populations of four- and five-year-olds. In our present state of knowledge we cannot specify, in advance, which children will benefit from such instruction and which will not, but the results clearly suggest that some will. In the case of the Transitivity Test, those five-year-olds who could use the transitive property had slightly higher scores in IQ and on the Verbal Maturity Test, but there is no case at all for attempting any instruction using similar populations with a view to improving the use of the transitive property before five years of age, and even at that age results are likely to be limited. The Conservation of Length Test Level II involves a different ability, and is more difficult than Level I. This is clearly in line with Smedslund's argument. Here, goal object and inference pattern are invariant, but percept changes. Again on the second administration of the test of reflexive and nonreflexive properties only 30 percent met the criterion, and of these only one child met the criterion on the Transitivity Test, although there were nine others who did meet the latter criterion. From this it may be deduced that the use of reflexive and nonreflexive properties, as measured, is neither a necessary nor a sufficient condition for the ability to use the transitive property of length relations. However, there are unavoidable changes between these tasks in percepts, and inference patterns. Consequently, the necessary change in percepts may have influenced the relations between the tasks, although this is not certain.

The paper by Douglas T. Owens is concerned with five- and six-year-old Negro children who were disadvantaged in the sense that they came from low-income families. No measures of IQ were given. A rather small number of children were subdivided into a Partial Treatment and Full Treatment group, with the former group receiving Instructional Units I and II, and the latter Instructional Units I, II, III, and IV. Unit III was designed to develop the ability of children to maintain relations between sets when the physical matching of objects was destroyed, while Unit IV was to help the children use the transitive property of matching relations. Transfer was inferred from a significant difference in favor of the full treatment group in performance on some property for which no instruction was given, provided that there was a significant change in the same direction on a related property for which instruction was given. On the posttest, the scores on the Transitivity of Matching Relations test improved significantly and scores on the test of the asymmetric property of matching relations improved significantly ($p < .01$). It very much looks as if the Instructional Units III and IV improved the performance of children on tasks which were similar to the activities involved in the treatment, but that there was no transfer to the Transitivity Problem, nor to the Transitivity of Length Relations Test. Again Unit III did not result in any improvement in the conservation ability of the Full Treatment Group. On the other hand, age was a factor influencing performance on all matching relations tests other than transitivity. But it had no effect on the abilities of children to use length relational properties.

We should not be too surprised that more children attained conservation of length relations than conservation of matching relations, and more children attained transitivity of length relations than transitivity of matching relations. In each case we have inference patterns constant but different goal objects and percepts. The source of the difference in difficulty between items on length and items on matching lies either in the nature of the goal object, or differences in percepts. While in Smedlund's study there was a tendency for conservation and transitivity of quantity to precede, respectively, conservation and transitivity of length, there were exceptions. With this rather small sample of somewhat limited background experience, these particular results can be accepted without too much surprise. Again, studies have generally shown that most children use conservation of a particular relational category before they use the transitive property for that category. In this study, however, about one-half of those using the transitive property in a particular category failed to conserve that category. As Owens suggests, these results may be to some extent interpreted in terms of the treatment effect. But we must not forget the erratic performance of dull and/or disadvantaged children. It seems that for them concrete reasoning has very narrow applicability for a much longer period compared with normal pupils, and that every situation, or nearly so, has to be tackled afresh. Put the other way around, there is little transfer of training. This erratic performance is also frequently seen in individual examinations using the various versions of the Binet Test of measured intelligence.

The educational implications of this study may be read as follows. For samples of similar children it is likely that instruction in the activities indicated is likely to improve performance only on closely related tasks. Such children must, of course, be given a full range of relevant experiences -- the opportunity to assimilate -- but teachers must not be disappointed in their slower progress and in their greater specificity with respect to performance, compared with pupils from more advantaged homes. There seems little evidence at present that the growth of logical thinking as such, in this type of pupil, will be aided by training.

We now turn to the study of David C. Johnson. In this study the children were drawn from kindergarten and from first grade. Only those children with an IQ between 80 and 120 were included in the study. Precise details of social class are not given. The aims of the investigation were (1) to determine if specific instruction improved the ability of the children both to form classes and establish selected equivalence and order relations, and (2) to see if transfer of training took place to certain other selected tasks. The five posttests measured respectively:

1. The ability to use the logical connectives "and," "or," and "not." (Test CA)

2. Understanding of the relations "more than," "fewer than," "as many as," "same shape as," and "same color as." (RA)

3. The ability to use two or more criteria at once. (MU)

4. The ability to solve class inclusion problems (CI). Success on this demands operatory classification in Piaget's view.

5. The ability to use the transitive property of the relations tested in the RA Test (TR).

From these tests five achievement and four transfer tasks were selected. For the sake of clarity these are listed in Table 3 for the benefit of readers, as it is otherwise difficult to hold in mind the differences between the achievement and transfer measures.

In considering the results it should be remembered, as was pointed out earlier, that for each age and IQ level there were only 10 children in the experimental and 10 in the control group. However, for all five achievement tests, the F values for Treatment and Intelligence were significant at the one percent level, although the mean score on MU_r remained low even in the experimental group. The F value for Age was not significant. In the case of the four transfer tests, the F value for Treatment was significant in the case of MU_3, MU_2, and TR, although the mean score

Table 3
Tests of the David C. Johnson Study

Achievement Tests	Content
CA_1	First ten items of CA
CA_2	Last ten items of CA (novel material)
CA_3	Intersection of Tests CA_1 and CA_2
RA	As RA
MU_r	Last six items of MU (intersection rings)
Transfer Tests	
MU_3	First six items of MU (3 x 3 matrices)
MU_2	Second six items of MU (2 x 2 matrices)
CI	Same as CI
TR	Same as TR

in the experimental group for MU_2 remained rather low. Against this, the F value in respect of intelligence was significant in the case of MU_2, CI, and TR. But age was not a significant variable.

Looking at the evidence as a whole it seems that, for similar populations, using the kinds of instructional activities undertaken in the study, help can be given in forming classes using intersection, union, and negation, and in making "prenumber" sets of objects. Moreover, there was some transfer to related activities. But Johnson rightly asks whether there was any real improvement in operativity. That cannot be answered for certain, since it depends on the criteria we decide on to define operativity, as was indicated earlier. As was pointed out above, the performance of the experimental and the control groups on MU_r remained low in spite of the fact that the former group did significantly better than the latter group. Moreover, as Johnson points out, the type of performance of children in the control group mitigates against interpreting it as an inability to form interesting rings. The treatment did not produce a significant improvement in the case of the CI Test, and although intelligence was a factor in performance on this test, neither the experimental nor the control group reached operatory classification. Inhelder and Piaget (1964, p. 164) bring evidence that graphic solutions to matrix items reach a maximum at age of six years, after which graphic solutions decline and operational solutions increase. It is, therefore, difficult to be sure, considering the ages of the pupils engaged in this study, that the improvement in the performance of the experimental group on the matrix items implied an improvement in logical thought. Again in the case of the transitivity test, the author suggests that improvement can be attributed to clarity of language rather than to the use of the transitivity property as such. Thus, looking at the data as a whole, it is possible that these children acquired physical knowledge and an increase in figurative knowledge, but not in logical-mathematical knowledge, except perhaps for those five-year-olds with high-measured IQ. Once again, it seems that the kinds of instruction given can be of considerable use to similar samples of children, but teachers must not think that it will necessarily bring about a growth in logical thought.

Johnson makes an important point when he advocates the need for a study of the actual relations that exist between the words "and," "or," and "not," and the growth of conjunction, disjunction, and negative concepts, respectively. It would also have been useful in this, and other studies, if a principal components analysis had been carried out in order to establish how the test scores cluster together and hence make possible some estimate of the abilities underlying the various tasks. This might have thrown light on what abilities the 3 x 3 and 2 x 2 matrix questions, and the intersecting rings questions, measure in these age groups.

A study concerned with the learning of classification and seriation is reported by Martin L. Johnson. Two groups of first- and second-grade children were involved. One was drawn from a Model Cities area and consisted of Negroes, and the other group came from a middle-class Caucasian neighborhood. No measured IQs are given. After the teaching of Unit I, pupils meeting the criterion test involving the relations "same length as," "shorter than," and "longer than," were randomly assigned to an experimental and control group. The members of the former group were given Instructional Unit II, which was designed for experiences in classifying objects on the basis of the equivalence relation "same length as," and in seriating on the basis of order relations "longer than," and "shorter than." The Conservation of Length Relations Test and Transitivity of Length Relations Test were used both as pretests and posttests, with the respective pretests used as covariates, whereas the Seriation and Classification Tests were used as posttests only.

Both grade and treatment significantly affected performance on the Seriation Test. In Item 1 of the Classification Test, both experimental and control groups did equally well, but in Item 2, where children had to discover the criteria for objects already classified, there was a slight relationship between treatment and improved performance ($.05 < p < .10$) where school and grade had no effect. Indeed, 75 percent of the subjects failed to achieve the criteria for classification. In the case of Item 3 (requiring the formation of a class with one element) neither school, grade, nor treatment was significantly related to performance, although more Grade 2 than Grade 1 pupils gave complete solutions. However, the fact that school and grade, but not treatment, affected posttest performance on the Conservation of Length Relations Test was, perhaps, somewhat surprising, and we shall come back to this point in a moment. Again, only school, and not treatment, affected the posttest performance of the Transitivity of Length Relations Test.

A number of issues are immediately raised by the results. First, the abilities involved in seriating sticks and strings are clearly related, but the seriation of the former is easier than the seriation of the latter. This is an example of percept and materials changing, with goal object and inference-pattern remaining constant. There is also a marked relation between the ability to insert a stick into an existing series of sticks with and without a baseline. And while the ability to seriate a set of "linear" objects can clearly be improved with training (the ability also improves with grade) in the case of both ethnic groups, we remain uncertain whether the level of logical thinking or the operativity of the pupils increased, or whether some rule was learned which enabled them to seriate more easily. The fact that no significant relation was found between transitivity of "longer than" and "shorter than" and the ability to seriate using the relations, might suggest that the training helped pupils to use an algorithm. On the other hand, it could be that more pupils had become operational in the Piagetian sense, but that the ability to seriate does not imply transitivity. Only a more precise study of the relationship between seriation and transitivity, keeping goal objects and percepts as constant as possible, will throw light on the latter problem.

If, however, the training did not improve operativity, then some of the other results are more understandable. For example, treatment did not affect performance on Items 2 and 3 of the Classification Test, on the Conservation of Length Relations Test, and on the Transitivity of Length Relations Test. Such results would be in keeping with the view that operativity was not increased by the treatment, as would the fact that performance on Items 1 and 2 of the Classification Test was only slightly related to the transitivity performance of "same length as," and performance on Item 3 not at all. On balance, it seems likely that the level of logical thinking was not increased, but grade did affect performance on some tests.

The study of R. A. Lesh concerns the interdependence of classification, seriation, and number concepts. Three parallel sequences of tasks (indicated by C1, C2 ---- C6: S1 ---- S6: N1 ------ N6) were devised to exhibit a relatively invariant sequential mastery, so that there would be only a small chance that a child could respond to the (n + 1)th task before he could respond to the nth task. The tasks were tried out by having them administered individually to each of 160 children aged 5 years 4 months to 6 years 7 months, living in a typical small town. This was in the nature of a pilot experiment leading up to a training study in which an answer was sought to the question of whether significant transfer of learning is possible between tasks which are characterized by isomorphic operational structures. In this case, the attempt was made to bring about the elaboration of number concepts through teaching seriation and classification concepts. But, in the training study, an answer to another question was sought: namely, would such training transfer to two tasks involving only spatial transformations which were roughly equivalent in difficulty to tasks N5 and N6 ?

The two spatial tasks chosen were Piaget's "three mountains" tasks (T6) and Piaget's tasks dealing with horizontal axes relative to the water level in an inclined bottle. We are not told how these particular tasks were made equivalent in difficulty to tasks N5 and N6, nor the exact details of materials and procedures used in these experiments. Experience shows that there are subtle changes in pupils' performance on these two tasks, depending on the precise task, on the materials used, and on the wording used in questioning (Lovell, 1972). Children are often at different stages on the two tasks, and, indeed, the level of performance may differ within a task because of changes in experimental procedure.

For the training study, another 20 children were chosen, aged 5 years 3 months to 6 years 2 months. These had correctly responded to tasks S1, N2, and C1, but had failed on tasks S3; N3, C3, and, from the experience gained in the pilot study, were most unlikely to respond correctly to tasks N5 and N6. The pupils were divided into an experimental and control group, with each individual in one group matched for sex and scores on certain other tests with an individual in the other group. Moreover, the primary purpose of the training was to get the pupils to coordinate schemes of actions which would lead to the groupings of seriation, multiple seriation, classification, and multiple classification; these, in turn, hopefully leading to increased decentration and analytic thinking.

The training successfully carried over to the number tasks but not to the spatial tasks, in the sense that the experimental group outperformed the control group on the former tasks, whereas none of the experimental group responded correctly to tasks T6 and T7. It would appear, as the author says, that training can enable some pupils to internalize schemes of actions which can be generalized to new situations involving the same schemes; in other pupils the transfer will be limited or nonexistent. The difficulty is that we do not know whether or not those in the experimental group who were able to pass tests N5 and N6 after the training were at some level of "transition" stage at the beginning. Basically we need to know more about the schemes available to children at the beginning of the training period than we are told in this and in all the other studies reported. But the author seems aware of the need to know more about the schemes at the outset. Indeed, he makes the point that the evidence is in favor of Piaget's distinction between three basic types of logical-mathematical operations being a useful one, and points out that children who spontaneously mastered all facts in the classification and seriation tasks of the pilot study outperformed the 10 children in the experimental group on number tasks. The schemes of the former were obviously different in some way from those of the latter. Lesh further points out that while the number concept involves only classification and seriation operations, the conservation of number also involves transformations. Clearly the schemes of some children in the experimental group permitted the handling of the relevant spatial transformation in the number field, while the schemes of other children did not. Such transformation had not been part of the training program. Yet, none of the pupils in the experimental group had schemes appropriate for the successful completion of tasks T6 and T7. Now, it is true that the scheme is a generalizable aspect of coordinating actions that can be applied to analogous situations, while schemes are coordinated among themselves in higher order structures. Moreover, tasks T6 and T7 depend for their successful completion on schemes and structures that enable them to handle aspects of projective and Euclidean space, respectively. Piaget is clear that the elaboration of logical-mathematical structures depends on maturation, social interchange, education and culture (cross cultural studies confirm this), and above all, on self-regulation. It would seem, then, that the structures involved for the successful completion on tasks T6 and T7 require different experiences from those required for N5 and N6: not that experience is a sufficient condition for the elaboration of structures, but that it is a necessary one. The discrepancy between performance on N5 and N6 on the one hand and T6 and T7 on the other, is not unexpected with this age group. The intra-individual variability in performing tasks characterized by a single operational structure is also likely to depend in part on experience and familiarity with materials, but to be sure we need to know much more about information processing in children than we know at present.

The author points to the rather narrow base of the training program and suggests that those pupils who spontaneously solved the classification and seriation tasks have better developed schemes than members of the experimental group, since the former did better than the latter on the number tests. This, he argues, is a case for more widely based teaching programs than many mathematics educators have admitted to hitherto. With this point the writer would heartily agree. We require widely based but directed programs, involving both action and language in small-group work. There seems little doubt that one can accelerate the performance of children on a particular task or on ones closely related to it, but the extent of transfer depends on the nature of the scheme available at the beginning of the teaching and on the width of zone of the pupil's assimilating capacity.

So far we have considered five studies. The effects of the training programs are briefly summarized in Table 4.

Table 4
Summary of the Training Programs

Group	Nature of Training	Effects of Training
1. Four- and five-year-olds. Normal spread of IQ and social background.	To establish length relations between two curves, to use reflexive and nonreflexive properties and to conserve length relations.	Improved ability to compare the lengths of two curves, in conservation of length relations, in use of reflexive and nonreflexive properties. Limited improvement in use of transitive property by 5-year-olds.
2. Five- and six-year-olds. Disadvantaged Negro children.	To establish length relations, to conserve matching relations, and to use the transitive property of matching relations.	Improved performance on transitivity of matching relations -- a task similar to activities involved in treatment. No transfer to other tasks.
3. Kindergarten and first-grade children with measured IQ 80-120. No precise details of social background.	To form classes, intersection and union of classes, complement of classes, relations between classes and between class elements.	Improved performance on all five direct achievement tests and on three of the transfer tests, although not on the test of Class Inclusion. Some doubt remains as to whether there is any improvement in regard to operativity.
4. First and second grade children, Negroes and middle-class Caucasian pupils. No IQ's given.	To classify on basis of equivalence relation "same length as," and seriate on basis of order relations "longer than," "shorter than."	Improved performance on Seriation Test. No improvement on Classification Test, Conservation of Length Relations Test or Transitivity Test.
5. Aged 5:3 to 6:2. Drawn from small Indiana community. A spread of ability.	To classify and seriate.	Improved performance on number tests but not on tasks involving spatial transformations.

In summary form, the overall picture can be quickly grasped. Suitable teaching programs aimed at improving children's understanding of certain mathematical ideas can profitably be undertaken with kindergarten, grade 1, and grade 2 children, providing such children are not slow learners and do not come from disadvantaged homes. With such populations, some children's performance will improve more than would have been the case without directed experiences, and in some instances, there may be some transfer of training. But it remains uncertain whether there will be any real improvement in operativity as the result of narrowly based experiences -- further, there remains the problem of the criteria used to define operativity. In the writer's view, the kinds of training included in these experiments could be incorporated into teaching programs in which small groups of children work through the various directed activities. It is, of course, important that children are not unduly pressed when the schemes available to them are far from those required for the tasks. If this is indeed the case, they are likely to assimilate the ideas with distortion, turn away in distaste, or have a tenuous grasp of the ideas in question. Readers will also have noticed that here and there in the papers, age, grade, and intelligence significantly affect the posttest results. This surely suggests that a longer period of varied experiences and better developed schemes, or earlier developed schemes in the case of higher intelligence, play a marked role in understanding mathematical ideas.

As against this, the one study reported among disadvantaged children suggests that teachers must not be disappointed when there is improvement in performance only on tasks very close to those that have been taught. Teachers must be prepared for limited improvement and limited transfer effect. In the case of school-educable retarded children, a number of studies have shown (Lovell, 1971a) that the intercorrelation coefficients calculated for performance on tasks administered individually on Piagetian lines are much lower than in the case of normal children. The schemes available to a child of chronological age and mental age seven years on a Binet type scale, must be different in ways we do not understand from one of chronological age ten and mental age seven. The fact that there is so much less transfer in dull and disadvantaged children does not imply that the kinds of teaching activities that these papers have discussed should be denied them. But teachers have to move carefully to maintain motivation and interest, be more sensitive to the capacity of such children to assimilate their experiences, and be ready to defer such activities for a few months. At the same time, our experience suggests that such children need more direction, sensibly applied in the teaching, than do abler and more advantaged children.

We must, of course, bear in mind that the posttests were given immediately after the training period ended. None of the studies reported giving a posttest, say, six months later. It is not possible to conjecture what the findings would have been on the latter occasion. This is a reason for long term follow-up studies of children, for which the writer argues a case later.

Three other points must be made.

1. In Almy's (1970) study, which involved a large number of second-grade children, it was found that performance was irregular across Piaget-type tasks. The results obtained in the present studies are consonant with the view that, even at 7-8 years of age, intellectual structures are still in a formative stage.

2. These studies have thrown no light on the question of the analytic set--or on the awareness on the part of the child that certain logical relations inhere within the situation. This demands a certain suspiciousness on the part of the child when faced with a task. We cannot tell from these studies whether the child is unable to elaborate the logical structures as necessary, or whether he can but fails to in the sense that he is not "switched on" to the implications of the situation.

3. It might appear from these studies that training in these types of activities was good in itself, for so often performance in closely related tasks improved although there was little transfer to other tasks. This may well be the case, but we cannot be sure. Young children appear to elaborate logical structures out of their interaction with their general environment, as they play and as they experiment with the world about them. It could be that direct activities, however skillfully and humanely applied to hold the interest of children, could nevertheless have a deleterious effect in the long run in the sense that we do not know how well such directed activities can be incorporated into the ongoing structures without, so to speak, any damage.

The sixth study, that of T. P. Carpenter, did not involve training. Rather, he was looking at the performance of 75 Grade 1 and 54 Grade 2 children on tasks involving the conservation and measurement of liquids. More specifically, Carpenter attempted to determine if young children responded differently to visual and numerical cues or simply to the last cue available. At the same time, the study was designed to find the role of equivalence and order relations in conservation and measurement problems. In order to reduce the number of tasks given to any one child, the three conservation problems, and both sets of problems (Nos. 2 and 3) in which quantities were measured with two different units, were administered to 61 of the children; the three relations being tested for all three sets of problems (Part A). The remaining 68 children were given all the measurement problems (Nos. 2, 3, 4, 5) with Equivalence and Nonequivalence II relations (Part B). We have no precise details of the social background of the children other than that they were drawn from a predominantly rural community in Wisconsin.

Some interesting data emerge. In Part A, 25 of the 61 children were correct on the Equivalence conservation item with a 95 percent probability that between 27 percent and 56 percent of the population would respond correctly to this item. These figures give readers some idea of the frequency of correct response to what has been the most usual type of conservation problem. But in the case of correct answers to all the tasks in Part A the reasons adduced fell almost completely under two main headings: (1) reference to the previous state, or (2) a compensating relationship between height and width, or between the number of units and unit size. Moreover, the evidence indicates that there are no significant differences either between the conservation and measurement problems or between the two types of measurement problems given. Furthermore, there are no significant differences in performance between Equivalence and Nonequivalence II relations for any of the problems given.

In Part B it is important to note the following three points: (1) the great difficulty in measurement with indistinguishably different units, (2) how much easier the tasks were which involved measurement of unequal-appearing intervals with the same unit, and (3) the ease of the items involving measurement with the same unit with apparent inequality. Indeed, there is a significant difference between each of the four types of measurement problems in Part B with respect to performance. Furthermore, there is no significant difference in respect of difficulty between Equivalence and Nonequivalence II relations, although in the case where unequal-appearing quantities are measured with the same unit, the difference in difficulty is significant at the p = .05 but not at the p = .01 level. Nearly all the reasons for the correct responses were based on reference to the previous state. The author argues with some reason that from a consideration of the performance on the four items given in both Part A and Part B, when quantities are measured with the same unit they are easier than corresponding conservation problems.

This study indicated that there is no difference in difficulty between conservation problems and corresponding measurement problems in which the distracting cues are numerical; and that misleading numerical cues can produce the same errors as misleading visual ones. The fact that these children found the tasks in which quantities were measured with the same unit into inequality so easy is good evidence that they do not respond only, or even primarily, on the basis of visual cues. Carpenter argues that the most significant factor in determining which cues young children attend to is the order in which the cues appear. The present writer would like to see more evidence for this, although to be fair to the author, he does clearly point out, as he must, that the order of cues is not the only factor found to affect responses.

The evidence also suggests that relations between quantities being compared does not affect performance, since there is no significant difference in performance between conservation and measurement problems employing Equivalence relations and corresponding problems employing Nonequivalence II relations -- here we have different percepts but the same goal object (quantity) and inference pattern (comparison of quantities). And apart from the situation in which it is impossible to identify the larger unit, Nonequivalence I problems are less difficult than the corresponding problems employing Equivalence and Nonequivalence II relations, since the correct relation between quantities could be found from the distracting cues by focusing on the appropriate dimension; e.g., focusing on the size rather than the number of units.

Carpenter points out that by the end of the first grade almost all children recognize the quantities are equal if they measure the same number of units, and quantity A is greater than quantity B if A measures more units than B. The writer's experience generally confirms these ages or grades, except for very dull children. It is, of course, true as the author asserts, that children in grades 1 and 2 do not have well-developed concepts of measurement nor are they able to measure accurately in all instances. In a more structured situation, such as Carpenter employed, the measurement cues were forced on the pupils, whereas Piaget employed less structured tasks in which children had to measure themselves in order to have any measurement cues to respond to. Thus, the beginning of understanding with respect to measurement came earlier in this study than in Piaget's experiments. The present writer has pointed out elsewhere (Lovell, 1971a) that the structure of a problem affects its difficulty for children. This is as true at the level of formal operational thought as at the age of onset of concrete operational thought. Moreover, this point does indicate that some kinds of well-structured problems can be introduced to Grade 1 children, and indeed in kindergarten, in the form of play with water and sand.

The author concludes that the major reason for most conservation and measurement failures lies in centering on a single dominant dimension, and the development of concepts of conservation and measurement can be thought of in terms of the increasing ability to decentre. He proposes that children pass through four stages, and these can usefully be compared with four steps proposed by Inhelder (1972), which were determined through learning experiments involving conservation and class inclusion. Unfortunately we do not know how consistent, in respect of a stage, a child was across all the tasks Carpenter undertook, any more than we know how consistent a child was in passing or failing a task he undertook. However, his proposals were:

1. The child responds on the basis of a single dominant dimension, which can be either visual or numerical depending on the task set.

2. The child is capable of changing from one dimension to another, but within any one task he tends to remain focused on one dimension. Sometimes, conservation results from the child keeping the earlier state in mind, but he is unable to consider both the earlier state prior to the transformation and the present state at one and the same time, and deciding which set of cues provides the right basis for comparison. Children at this stage seem to ignore the present state and refer only to the former state ("they were the same before"). The author suggests that it is children at this transition stage who gain most from training in conservation, but such gains are on a narrow front and without a basis for movement to a later stage and hence to an improvement in operativity. Certainly the Geneva school would claim that training is ineffective except for those at a transition stage.

3. The schemes now permit the child to consider a number of conditions of a quantity, simultaneously, and choose the one that provides a rational basis for a comparison.

4. The child can now use the information from the prior and present state and find the correct relation between the sizes of different units based on the number of units which the two comparable quantities measure.

Inhelder's paper gives details of proposed stages found in learning experiments involving conservation and class inclusion:

1. Two different systems of evaluation (e.g., number of sticks and lengths of sticks) could be elicited, but neither was sufficiently developed to permit their integration. The two separate systems were activated successively, and the child did not feel there was any contradiction.

2. In place of the two evaluative schemes being evoked successively, both seemed to be present almost simultaneously. But the pupil could not conceive of a new solution which could take both schemes into account. However, he is conscious of contradictions.

3. An attempt is now made at integrating the two evaluative schemes, but it is inadequate. There are compromise solutions in the form of partial compensations.

4. The different schemes can now be integrated. Thus in respect of two lines of equal length composed of matches of different lengths the reply might be "You have more matches but they are shorter.". A scheme no longer operates a <u>post hoc</u> correction on another but rather there is a reciprocal adjustment.

Inhelder states that in all the processes that have come to light in the training experiments carried out at Geneva, development takes place in a similar way except in one group of problems. In those involving logical operations in the strict sense of the word (e.g., class inclusion), the regulatory mechanisms found at stage (2) which yield the awareness of contradictions are not followed by the compromise solutions of stage (3) but by complete logical compensations which later result in correct solutions (4).

The suggestions of Carpenter, also of the Geneva school, have been mentioned since the importance of studying the growth of partial structures is raised in the next section.

Some Suggestions for Further Research

The studies that have been presented have been carefully designed and skillfully executed, and the suggestions which follow are in no sense to be taken as a reflection on them, but rather they should be regarded as lines for research in the future. The research reported has been concerned with the period of four to seven or eight years of age -- a period in which Piaget characterized the child's thinking as a semi-logic, or a one-way mapping, and the suggestions made here necessarily relate to the same period and the years that immediately follow. Unfortunately, psychology has not yet provided a valid theory of cognitive development which is presented in detailed process terms. Piaget has given us great insights and his developmental theory of intellectual growth provides a useful conceptual framework in which the teacher can consider the ideas he wishes his pupils to develop, although it still leaves large gaps in our knowledge. However, his theory is constantly developing as new problems are raised, new methods developed to deal with these, and existing models adjusted or refined to account for these findings. The following broad research areas are suggested.

1. Many studies have enumerated the items which a child passes or fails. This certainly has its place either in training, or in the frequent monitoring of children in a longitudinal study. But perhaps at this moment of time more emphasis should be placed on carefully recording the precise nature of the response both in respect of action and verbalization. We need far more knowledge about the exact stage of development of the relevant schemes of a child at the beginning of the training. It would appear from the Geneva evidence that children at the lowest operative levels get little from the training. But when items are scored on a pass or fail basis we do not see the detailed base line, or detailed development of the schemes which lead to a correct solution to the problem. Such information is greatly needed.

Again when training programs employ two or more procedures (e.g., numerical [number of matches] and ordinal [length of matches] ways of dealing with problems of constructing or judging lengths) we need to see the interaction between schemes at various points or steps. At some point conflict ensues, and in the view of the Geneva school it is conflict that triggers the reciprocal assimilation between schemes which provides the final resolution of the problem. Note carefully, however, that the different schemes which are assimilated and integrated may not all be at the same developmental level. Much research is required here as, we know little about the effect on each other of the various activities in which the child engages.

Such studies are likely to throw light on the nature of the schemes (in respect of mathematical ideas) available to normal as compared with dull and disadvantaged pupils. It would also be likely to shed light on the important problem of transfer. In short, such studies are likely to throw light on processes that impel the pupil forward. The classical Piagetian structural model must be supplemented. We need more information on the growth of schemes underpinning mathematical ideas and the way in which they become integrated with other schemes.

In keeping with what has just been said, readers are reminded again of the views of Smedslund regarding the importance of studying the growth of inference patterns holding constant percept (as far as possible) and goal object and also of the point made by Pinard and Laurendeau (1969) that until we know more than we do now, the *structure d'ensemble* criterion should be investigated to see to what extent the different Piagetian groupings are achieved in synchrony on tasks related to the same conceptual content and same material. Hamel and van der Veer (1972) attempted to carry out the suggestion of Pinard and Laurendeau involving Multiple Classification and Multiple Seriation tasks. Using their method of scoring, the correlation between performance on these tasks was around +0.6 even when measured intelligence has been partialled out. The authors are well aware that some children may solve this kind of problem by means other than using operational schemes -- as we saw earlier. However, they also make two points. First, the amount of information affects problem-solving behavior at the stage of concrete operational thought. Second, there is a need for a longitudinal study of individual children as they move from the preoperational to the concrete operational stage of thought, to see to what extent different operatory schemes develop in synchrony, and how performance is affected by irrelevant variables. The views of Hamel and van der Veer strongly support the general views just proposed by the writer.

2. In connection with what has been just said it seems to the writer that an information processing model of some of the tasks that have been attempted in these papers could be of importance. It is true that at the Université de Montréal work in this field has been in progress for some time, although the results are not widely known yet. In order to encourage readers who have easy access to a computer, a brief sketch of a little of the Montreal work is given.(See Baylor et al. 1973.) In, say, weight seriation and length seriation tasks, a video-tape record is made of the child solving, or attempting to solve, the problem given. The actions and words of the child are then transcribed onto a protocol. The protocol is carefully analyzed, in respect of both the task environment and of the subject's intellectual structures as revealed by his behaviors. The pupil's behavior on the task is then simulated by writing a set of rules or program which can be interpreted by a digital computer. When the program is executed a close reproduction of the child's behavior is obtained. It can also generate further protocols on new but similar problems.

Baylor gives an example which related to the weight seriation task. The child is presented with seven white two-inch cubes and a sensitive balance. The former all looked alike but were identified by different letter names. Their weights varied from 100.2 gms to 106.5 gms so that the pupil was forced to use the balance to judge the relative weights of the cubes. He was not allowed to put more than two cubes on the scale at any one time. Examples are given of the recognizable strategies found in the task and programs are given for the various "stages" of strategies employed, namely:

1. A juxtaposition of pairs without any coordination between pairs;

2. The child tries to coordinate his successive weighings of pairs;

3. Successive weighings are coordinated.

The derivation of predictions from the model forms the basis for further experimental studies. Data from the latter may, of course, demand modifications in the model; in the way one could come to a number of recyclings through the series. For example, if a pupil at stage (3) was presented with two extra cubes and asked to intercalate these into the series, would he behave as the model predicts? If he does not, then why not? The model would have to be appropriately amended in this case.

At the time of writing, Baylor et al had a program which would simulate behavior over both a task which involved the seriation of sticks (taking only two at a time) and the weight seriation task, for a program must ever be made more general. At the same time these writers point out that there are certain aspects of Piagetian theory that have not yet found adequate nontrivial representations in the information processing models (e.g., such competences as the onset of reversible, asymmetric operations, and the ability to envisage, say, a cube as weighing more than its neighbor to the right and at the same time weighing less than its neighbor to the left). Thus, at present, an information-processing model and a Piagetian model must remain complementary, although the former may, as the result of future research, give operational definitions to the Piagetian insights that cannot be represented at present.

At the moment a small amount of work has been carried out in England involving information processing in respect of the class inclusion problem and the conservation of quantity, but details of the work are not yet available. In the writer's view this is a potentially useful research area as it may well throw great light on the nature of the partial structures; that is, schemes in the process of development, and of the interaction between schemes when a complete solution is developing. It may also throw light on the partial structures of able and less able children, and on the question of transfer.

3. More research is required into the growth of schemes leading to success in the class inclusion problem. Piaget believes that successful performance in this problem demands operatory classification, while Inhelder (1972) points out that the experience at Geneva suggests that training in the class inclusion problem has positive effects on performance on the conservation problems. In D. C. Johnson's study, children were given experience in forming classes, in forming intersections and unions, and in forming the complement of a class. The Class Inclusion test administered as a posttest involved the following factors: presence of an extraneous object, three or more subsets present, equal numbers in a set and its elements, mingled items, items not visually present, addition or subtraction of an item after an initial comparison. Now it is true that whenever a class and its complement are specified, the idea of inclusion is implicit. But apparently the training was of little avail when it came to the CI items. Further detailed studies involving the growth of schemes relating to the CI problem, taking into account variables such as Johnson used, are likely to be of value. They would also throw further light on the growth of schemes in relation to the amount of relevant and irrelevant information provided in the problems. And it might develop that greater transfer to other tasks such as conservation might be found, as Inhelder suggests.

Beilin reviews the relevant evidence concerning training studies in class inclusion. After reviewing the evidence provided by Kohnstamm and the counter-evidence of the Geneva school, together with other studies, he concludes that training can lead to the successful conclusion of this logical ability. Whether operative achievement from instruction and training results is still unresolved, since conceptual and operational definition of operativity have yet to be made. This need not detract the mathematics educator. He wishes to know whether the "pay off" in respect of transfer to other skills is greater if the emphasis is put more on class inclusion.

4. While the present papers involve fundamental logical structures vital to mathematics, which are elaborated with greater or lesser understanding in the move from semi-logic to concrete logical thought, we should not be content with these alone. We need to establish, in a whole range of abilities from very able to school-educable retarded, more knowledge as to the stages through which pupils pass in elaborating the concepts involved in, say, the numeration system or in the properties of the natural number system. And we need far more information in respect of the growth of spatial and geometrical concepts and how these develop in comparison with numerical ones. Again, we need more information on whether topological concepts develop prior to Euclidean and projective concepts in the child's representation of space. Lovell (1959), Lunzer (1960a), and Martin (1973) have doubted this. But the evidence of Laurendau and Pinard (1970) must also be considered. The outcome of this argument is important for it would give guidance in respect of the order in which spatial ideas should be introduced to young children.

5. Research is needed along the lines argued by Steffe (1973). In his paper he has attempted to outline some possible relations between the cognitive systems of the child and the mathematical systems of finite cardinal and ordinal number. Moreover, he argues that it seems likely that some mathematical structures may be more parsimonious models of cognitive operations than the genetic structures proposed by Piaget. For example, he reasons a case for suggesting that, on the basis of information which we have at present, the structures of connected, asymmetrical, transitive relations is more parsimonious as a logical model of seriation than is Piaget's Grouping V (Addition of Asymmetrical Relations). But more research is needed to establish if this is the case.

Again, since logical identity (essentially an equivalence relation) is an integral part of the Piagetian grouping structures, a great deal more knowledge needs to be established concerning its development, and of the relation between its development and the growth of seriation and classification behaviors. Moreover, although logical identity, set equivalence, and set similarity are all equivalence relations, they may have different roles in the child's elaboration of the concepts of cardinal and ordinal number. Here, too, research is needed, and Steffe lists a number of problems to be investigated holding in mind Piaget's Grouping I (Primary Addition of Classes).

6. Some of the basic research which we have suggested will require time both to execute and implement. Meanwhile children have to be educated. Thus, I also believe that a longitudinal study of children is important in which the style of teaching is kept constant over a period of years even though the same teachers do not remain with the pupils throughout the period. This involves great problems. For example, even if the style

of teaching remains constant over a number of years, the vigour and enthusiasm of individual teachers may not. In my view we need to compare the performance and understanding of pupils towards the end of the elementary school, whose teaching can be characterized since kindergarten and Grade 1 as carefully directed activities of the kinds indicated in these papers, with pupils who have been taught on more traditional lines. This is a formidable task. For one thing, the directed activities to induce an understanding of mathematical ideas would have to be done in a manner to hold the interest of pupils since children in the elementary school look upon mathematics as a tool with which to solve real-life problems, and not as a purely intellectual exercise. Again, the ability to compute quickly and accurately is a skill that all pupils need. Bluntly, our research to date has been too short-termed. In respect of the present studies, for example, it is unlikely that we shall ever know what happened to the mathematical understanding of these people in later school life. Do any improvements brought about by such directed activities enable the pupil to elaborate the concept of time any more easily, or have a surer grasp of the properties of the natural number system in the sixth grade? The design and execution of such a study would present formidable problems, but until it is tackled we shall not have accurate data on the long-term effects of a consistent style of teaching mathematics over a number of years.

The writer is, of course, well aware of the results of the Almy (1970) study, in which a large number of children who had had prescribed mathematics and science programs since kindergarten (AAAS, GCMP, and SCIS) were compared with pupils whose mathematics and science activities were planned by the teacher. In the second grade both groups of children did about as well on Piaget-derived tasks; so there was no indication of superior logical thought in either group. However, two points need to be made in respect of the Almy study. First, we cannot be sure how well the presribed programs were implemented; that is, of the quality of the teaching. Attention was drawn to this danger in the opening paragraphs of this chapter. In any study the writer had in mind the quality and style of teaching would be controlled as far as possible. If this were not possible from kindergarten to grade 6, it might be possible from kindergarten to grades 3 or 4. Second, the follow-up study of Almy was in the second grade; the writer has in mind a much longer study. It is, of course, unlikely that we should ever be able to maintain a style of teaching across all areas of school life, and we have no idea what effect on the growth of logical thinking the mathematics programs alone would have.

While mathematics educators must necessarily believe in the particular programs they advocate, we shall not know the long term effects of such programs until a well-designed experiment is undertaken. Beilin, in his extensive review of the literature in the training and acquisition of logical structures, argues that since the child can construct a conceptual system out of many materials and techniques, even those not intended by the researcher, the basic pattern of organizations is internal. That is to say, the growth of the logical operational system is under the control of a genetic mechanism and permits the growth of intellectual structures through interactions of environmental inputs. Put simply, we want to know what the long-term effects are of particular environmental inputs on the growth of logical thought and the grasp of mathematical ideas. And can young pupils assimilate particular inputs, as in the case of directed activities in mathematics, without any ill effects?

REFERENCES

Adler, M. J. Some educational implications of the theories of Jean Piaget and J. S. Bruner. *Canadian Educational and Research Digest*, 1964, 4, 291-305.

Almy, M., Chittenden, E., & Miller, P. *Young children's thinking: Studies of some aspects of Piaget's theory.* New York: Teachers College Press, 1966.

Almy, M. *Logical thinking in second grade.* New York: Teachers College Press, 1970.

Applebaum, M., & Bargmann, R. E. A fortran II program for MUDAID: Multivariate, univariate, and discriminant analysis of irregular data. *NONR 1834*, 39, Urbana: University of Illinois Press, 1967.

Baylor, W., Gascon, J., Lemoyne, G., & Pothier, N. An information processing model of some seriation tasks. *Canadian Psychologist*, 1973, 14, 167-196.

Bailey, T. On the measurement of polygonal paths by young children. Unpublished doctoral dissertation, University of Georgia, 1973.

Bearison, D. J. Role of measurement operations in the acquisition of conservation. *Developmental Psychology*, 1969, 1, 653-660.

Beilin, H. Perceptual-cognitive conflict in the development of an invariant area concept. *Journal of Experimental Child Psychology*, 1964, 1, 208-226.

Beilin, H. Learning and operational convergence in logical thought development. *Journal of Experimental Child Psychology*, 1965, 2, 317-339.

Beilin, H. Cognitive capacities of young children: A replication. *Science*, 1968, 162, 920-921.

Beilin, H. The training and acquisition of logical operations. In M. F. Rosskopf, L. P. Steffe, & S. Taback (Eds.), *Piagetian cognitive-development research and mathematical education.* Washington, D.C.: National Council of Teachers of Mathematics, 1971.

Beilin, H., & Franklin, I. C. Logical operations in area and length measurement: Age and training effects. *Child Development*, 1962, 33, 607-618.

Berlyne, D. E. *Structure and direction in thinking.* New York: John Wiley & Sons, 1965.

Beth, E. W., & Piaget, J. *Mathematical epistemology and psychology.* Dordrecht-Holland: D. Reidel, 1966.

Birkhoff, G., & Maclane, S. *A survey of modern algebra.* New York: Macmillan, 1958.

Bourbaki, N. L'architecture des mathématiques. *Le Lionnais,* Paris, 1948, 35-37.

Braine, M. D. S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. *Psychological Monographs: General and Applied,* 1959, 73 (5, Whole No. 475).

Brownell, W. A., & Chazal, C. B. The effects of premature drill in third-grade arithmetic. *Journal of Educational Research,* 1935, 29, 17-28.

Bruner, J. S. Inhelder and Piaget's *The growth of logical thinking:* I. A psychologist's viewpoint. *British Journal of Psychology.,* 1959, 50 363-370.

Bruner, J. S. The course of cognitive growth. *American Psychologist,* 1964, 19, 1-15. (a)

Bruner, J. S. Some theorems of instruction illustrated with reference to mathematics. In E. R. Hilgard (Ed.), *Theories of learning and instruction.* Sixty-third Yearbook of the National Society for the Study of Education, 1964, 306-336. (b)

Bruner, J. S. *Toward a theory of instruction.* New York: Norton, 1968.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. *A study of thinking.* New York: John Wiley & Sons, 1956.

Bruner, J. S., & Kenney, H. J. Representation and mathematics learning. In L. N. Morrisett & J. Vinsonhaler (Eds.), *Mathematical Learning, Monographs of the Society for Research in Child Development,* 1964, 30 (1, Serial No. 99).

Bruner, J. S., & Kenney, H. J. On multiple ordering. In J. S. Bruner, R. R. Olver, & P. M. Greenfield, et al., *Studies in cognitive growth.* New York: John Wiley & Sons, 1966.

Bruner, J. S., Olver, R. R., & Greenfield, P. M., et al. *Studies in cognitive growth.* New York: John Wiley & Sons, 1966.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching.* Chicago: Rand McNally, 1963.

Carey, R., & Steffe, L. P. An investigation in the learning of equivalence and order relations by four-and-five-year-old children. *Research Paper No. 17, Georgia Research and Development Center in Educational Stimulation.* Athens: University of Georgia, 1968.

Carpenter, T. P. The role of equivalence and order relations in the development and coordination of the concepts of unit size and number of units in selected conservation type measurement problems. *Technical Report No. 178, Wisconsin Research and Development Center for Cognitive Learning.* Madison: The University of Wisconsin, 1971. (a)

Carpenter, T. P. The performance of first grade students on a nonstandard set of measurement tasks. *Technical Report No. 211, Wisconsin Research and Development Center for Cognitive Learning*. Madison: The University of Wisconsin, 1971. (b)

Churchill, E. M. The number concepts of the young child: Part I. *Researches and Studies*, 1958, 17, 34-49. (a)

Churchill, E. M. The number concepts of the young child: Part II. *Researches and Studies*, 1958, 18, 28-46. (b)

Clarke, A. M., Cooper, G. M., & Loudon, E. H. A set to establish equivalence relations in pre-school children. *Journal of Experimental Child Psychology*, 1969, 8, 180-189.

Coxford, A., Jr. The effects of instruction on the stage placement of children in Piaget's seriation experiments. *Arithmetic Teacher*, 1964, 11, 4-9.

Cochran, W. G. The comparison of percentages in matched samples. *Biometrika*, 1950, 37, 256-266.

Darnell, C. D., & Bourne, L., Jr. Effects of age, verbal ability, and pre-training with component concepts on the performance of children in a bidimensional classification task. *Journal of Educational Psychology*, 1970, 61, 66-71.

Dienes, Z. P. *Building of mathematics*. (Rev. ed.). London: Hutchinson, 1971.

Divers, B. P., Jr. The ability of kindergarten and first grade children to use the transitive property of three length relations in three perceptual situations. Unpublished doctoral dissertation, University of Georgia, 1970.

Dodwell, P. C. Children's understanding of number and related concepts. *Canadian Journal of Psychology*. 1960, 14, 191-205.

Dodwell, P. C. Children's understanding of number concepts: Characteristics of an individual and of a group test. *Canadian Journal of Psychology*, 1961, 15, 29-36.

Dodwell, P. C. Relations between understanding of the logic of classes and of cardinal number in children. *Canadian Journal of Psychology*, 1962, 16, 152-160.

Edwards, J. Effects of instruction and concomitant variables on multiple categorization ability. *Journal of Educational Psychology*, 1969, 60, 138-143.

Elkind, D. Piaget's conservation problems. *Child Development*, 1967, 38, 841-848.

Elkind, D. Discrimination, seriation, and numeration of size and dimensional differences in young children: Piaget replication study VI. *Journal of Genetic Psychology*, 1964, 104, 275-296.

El'konin, D. B., & Davydov, V. V. Children's capacity for learning mathematics. L. P. Steffe (Ed.), *Soviet studies in the psychology of learning and teaching mathematics,* Vol. 7, 261, School Mathematics Study Group, 1974.

Feller, W. *An introduction to probability theory and its applications.* New York: John Wiley & Sons, 1957.

Ferguson, G. A. *Statistical analysis in psychology and education.* New York: McGraw-Hill, 1966.

Finn, J. D. *Multivariance. Fortran program for univariate and multivariate analysis of variance and covariance.* Buffalo, State University of New York, 1967.

Flavell, J. H. *The developmental psychology of Jean Piaget.* Princeton: D. Van Nostrand Company, 1963.

Fleischmann, B., Gilmore, S., & Ginsburg, H. The strength of nonconservation. *Journal of Experimental Child Psychology,* 1966, 4, 353-368.

Gal'perin, P. Ya., & Georgiev, L. S. The formation of elementary mathematical notions. In J. Kilpatrick & I. Wirszup (Eds.), *Soviet studies in the psychology of learning and teaching mathematics,* Vol. 1. School Mathematics Study Group, 1969.

Gast, H. Der Umgang mit Zahlen und Zahlbegilden in der frühen Kindheit. Z. Psycholo., 1957, 161, 1-90. Cited in I. E. Sigel & F. H. Hooper (Eds.), *Logical thinking in children.* New York: Holt, Rinehart, & Winston, 1968, p. 97.

Gödel, K. Über formal unetscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Mathematiker Zeitschrift,* 1934, 39, 176-210, 405-431.

Greco, P., Grize, J., Papert, S., & Piaget, J. Problems de la construction du nombre. *Études d'épistémologie génétique.* Vol. 2. Paris: Presses Université de France, 1960.

Greenhouse, S. W., & Geisser, S. On methods in the analysis of profile data. *Psychometrika,* 1959, 24, 95-112.

Griffiths, J. A., Shantz, C. A., & Sigel, I. E. A methodological problem in conservation studies: The use of relational terms. *Child Development,* 1967, 38, 841-848.

Grize, J. B. Du groupment au nombre. In P. Greco, J. B. Grize, S. Papert, & J. Piaget, Problems de la construction du nombre. *Études d'épistémologie génétique.* Paris: Presses Université de France, 1960, 11, 69-96.

Guilford, J. P. *Fundamental statistics in psychology and education.* New York: McGraw-Hill, 1956.

Hamel, B. R., & van der Veer, M. A. A. Structure d'ensemble, multiple classification, multiple seriation and amount of irrelevant information. *British Journal of Educational Psychology*, 1972, 42, 319-325.

Harper, E. H., & Steffe, L. P. The effects of selected experiences on the ability of kindergarten and first grade children to conserve numerousness. *Technical Report No. 38, Wisconsin Research and Development Center for Cognitive Learning.* Madison: The University of Wisconsin, 1968.

Herstein, I. N. *Topics in algebra.* Waltham: Blaisdell Publishing Company, 1964.

Hilgard, E. R. A perspective on the relationship between learning theory and educational practices. In E. R. Hilgard (Ed.), *Theories of learning and instruction.* Sixty-third Yearbook of the National Society for the Study of Education, 1964, 402-415.

Holowinsky, Ivan. Seriation actions in preschool children. *Journal of Learning Disabilities,* 1970, 9, 34-35.

Hooper, F. Piaget's conservation tasks: The logical and developmental priority of identity conservation. *Journal of Experimental Child Psychology,* 1969, 8, 234-249.

Huntington, J. R. Linear measurement in the primary grades: A comparison of Piaget's description of the child's spontaneous conceptual development and the SMSG sequence of instruction. *Journal for Research in Mathematics Education,* 1970, 1, 219-232.

Hyde, D. M. An investigation of Piaget's theories of the development of the concept of number. Unpublished doctoral dissertation, University of London, 1959.

Inhelder, B. Information processing tendencies -- empirical studies. In Farnham - Diggory, S. (Ed.), *Information processing in children.* New York: Academic Press, 1972.

Inhelder, B., & Piaget, J. *The growth of logical thinking from childhood to adolescence.* Translated by A. Parsons & S. Milgram, New York: Basic Books, 1958.

Inhelder, B., & Piaget, J. *The early growth of logic in the child: Classification and seriation.* Translated by E. A. Lunzer, London: Routledge and Paul, 1964.

Inhelder, B., & Sinclair, H. Learning cognitive structures. In P. Mussen, J. Langer, & M. Covington (Eds.), *Trends and Issues in Developmental Psychology.* New York: Holt, Rinehart, & Winston, 1969.

Johnson, D. C. An investigation in the learning of selected parts of a boolean algebra by five- and six-year-old children. Unpublished doctoral dissertation, University of Georgia, 1971.

Johnson, M. L. Effects of selected experiences on the classification and seriation abilities of young children. Unpublished doctoral dissertation, University of Georgia, 1971.

Kilpatrick, J., & Wirszup, I. (Eds.). Preface. *Soviet studies in the psychology of learning and teaching mathematics,* Vol. II. School Mathematics Study Group, 1969.

Kohnstamm, G. A. *Piaget's analysis of class inclusion: Right or wrong?* Translated by M. Reck-O'Toole & G. Uildriks-Bone, The Hague: Mouton & Co., 1967.

Lasry, J. C., & Laurendeau, M. Apprentissage empirique de la notion d'inclusion. *Human Development,* 1969, 12, 141-153.

Laurendeau, M., & Pinard, A. *The development of the concept of space in the child.* New York: International Universities Press, 1970.

Lesh, R. A. The generalization of Piagetian operations as it relates to the hypothesized functional interdependence between class, series, and number concepts. Unpublished doctoral dissertation, Indiana University, 1971.

Lovell, K. A follow-up study of some aspects of the work of Piaget and Inhelder on the child's conception of space. *British Journal of Educational Psychology,* 1959, 29, 104-117.

Lovell, K. Concepts in mathematics. In H. J. Klausmeier & C. W. Harris (Eds.), *Analyses of concept learning.* New York: Academic Press, 1966. (a)

Lovell, K. In The developmental approach of Jean Piaget: Open discussion. In M. Garrison, Jr. (Ed.), *Cognitive models and development in mental retardation. Monograph supplement to the American Journal of Mental Deficiency,* 1966, 70 (4), 84-89. (b)

Lovell, K. The development of some mathematical ideas in elementary school pupils. In M. F. Rosskopf, L. P. Steffe, & S. Taback (Eds.), *Piagetian cognitive-development research and mathematics education.* Washington: National Council of Teachers of Mathematics, 1971. (a)

Lovell, K. *The growth of understanding in mathematics: Kindergarten through grade three.* New York: Holt, Rinehart & Winston, 1971. (b)

Lovell, K. Intellectual growth and understanding in mathematics. Paper presented at the meeting of the American Educational Research Association, New York, 1971. (c)

Lovell, K. Intellectual growth and understanding mathematics. *Journal for Research in Mathematics Education,* 1972, 3, 164-182.

Lovell, K., Mitchell, B., & Everett, I. R. An experimental study of the growth of some structures. *British Journal of Psychology,* 1962, 53, 175-188.

Lovell, K., & Ogilvie, E. The growth of the concept of volume in junior high school children. *Journal of Child Psychology and Psychiatry*, 1961, 2, 118-126.

Lovell, K., Healey, D., & Rowland, A. D. Growth of some geometrical concepts. *Child Development*, 1962, 33, 751-767.

Lunzer, E. A. *Recent studies in Britain based on the work of Jean Piaget*. London: National Foundation for Educational Research in England and Wales, Occasional publication, 4, 1960. (a)

Lunzer, E. A. Some points of Piagetian theory in the light of experimental criticism. *Journal of Child Psychology and Psychiatry*, 1960, 1, 191-202. (b)

Martin, J. L. An investigation of the development of selected topological properties in the representational space of young children. Paper presented at the meeting of The American Educational Research Association, New Orleans, 1973.

McLellan, J. A., & Dewey, J. *The psychology of number*. New York: D. Appleton, 1914.

McManis, D. L. Conservation of identity and equivalence of quantity by retardates. *Journal of Genetic Psychology*, 1969, 115, 63-69.

Maccoby, M., & Modiano, N. On culture and equivalence: I. In J. S. Bruner, R. R. Olver, & P. M. Greenfield, et al., *Studies in cognitive growth*. New York: John Wiley & Sons, 1966.

Northman, J. E., & Gruen, G. E. Relationship between identity and equivalence conservation. *Developmental Psychology*, 1970, 2, 311.

Olmsted, P., Parks, C. V., & Rickel, A. The development of classification skills in the preschool child. *International Review of Education*, 1970, 16, 67-80.

Olver, R. R., & Hornsby, J. R. On equivalence. In J. S. Bruner, R. R. Olver, & P. M. Greenfield, et al., *Studies in cognitive growth*. New York: John Wiley & Sons, 1966.

Ostle, B. *Statistics in research*. Ames: Iowa State University Press, 1963.

Otis, A. S., & Lennon, R. T. *Manual for administration: Otis-Lennon mental ability test*. New York: Harcourt, Brace & World, 1967.

Owens, D. T. The effects of selected experiences on the ability of disadvantaged kindergarten and first grade children to use properties of equivalence and order relations. Unpublished doctoral dissertation, University of Georgia, 1972.

Owens, D. T., & Steffe, L. P. Performance of kindergarten children on transitivity of three matching relations. *Journal for Research in Mathematics Education*, 1972, 3, 141-154.

Piaget, J. Classes, relations et nombres: *Essai sur les "groupements" de la logistique et la réversibilité de la pensée.* Paris: Vrin, 1942.

Piaget, J. *The child's conception of number.* London: Routledge and Kegan Paul, 1952.

Piaget, J. Equilibration and the development of logical structures. In J. M. Tanner & B. Inhelder (Eds.), *Discussions on child development.* Vol. 4. London: Tavistock, 1960.

Piaget, J. Development and Learning. In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered.* A report of the conference on cognitive studies and curriculum development. Ithaca, New York: Cornell University Press, 1964. (a)

Piaget, J. Mother structures and the notion of number. In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered.* A report of the conference on cognitive studies and curriculum development. Ithaca, New York: Cornell University Press, 1964. (b)

Piaget, J. *The psychology of intelligence.* London: Routledge and Kegan Paul, 1964. (c)

Piaget, J. Quantification, conservation, and nativism. *Science,* 1968, 162, 976-979.

Piaget, J. *Genetic epistemology.* New York: Teachers College Press, 1970.

Piaget, J. *Science of education and the psychology of the child.* Translated by D. Coltman, New York: Viking, 1971.

Piaget, J. Forward. In M. Schwebel & J. Raph (Eds.), *Piaget in the classroom.* New York: Basic Books, Inc., 1973. (a)

Piaget, J. *To understand is to invent: The future of education,* New York: Grossman, 1973. (b)

Piaget, J., & Inhelder, B. *The child's conception of space.* Translated by F. J. Langdon and J. L. Lunzer. New York: Norton, 1967.

Piaget, J., & Inhelder, B. *The mental imagery of the child.* Translated by P. A. Chilton. New York: Basic Books, Inc., 1971.

Piaget, J., Inhelder, B., & Szeminska, A. *The child's conception of geometry.* New York: Harper and Row, 1960.

Pinard, A., & Laurendeau, M. "Stage" in Piaget's cognitive development theory: Exegesis of a concept. In D. Elkind & J. H. Flavell (Eds.), *Studies in cognitive development.* New York: Oxford University Press, 1969.

Ripple, R. E., & Rockcastle, V. N. (Eds.). *Piaget rediscovered.* A report of the conference on cognitive studies and curriculum development. Ithaca, New York: Cornell University Press, 1964.

Rothenberg, B. B. Conservation of number among four-and-five-year-old children: Some methodological considerations. *Child Development*, 1969, 40, 383-406.

Roughhead, W. G., & Scandura, J. M. What is learned in mathematical discovery. *Journal of Educational Psychology*, 1968, 59, 282-289.

Saltz, E., & Sigel, I. E. Concept overdiscrimination in children. *Journal of Experimental Psychology*, 1967, 73, 1-8.

Seigel, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill Book Company, 1956.

Shantz, C. U. A developmental study of Piaget's theory of logical multiplication. *The Merrill-Palmer Quarterly*, 1967, 13, 121-137.

Siegel, L. S., & Goldstein, A. G. Conservation of number in young children: Recency versus relational response strategies. *Developmental Psychology*, 1969, 1, 128-130.

Sigel, I. E., & Hooper, F. H. (Eds.). *Logical thinking in children.* New York: Holt, Rinehart and Winston, 1968.

Sigel, I. E., Roeper, A., & Hooper, F. H. A training procedure for acquisition of Piaget's conservation of quantity: A pilot study and its replication. *British Journal of Educational Psychology*, 1966, 36, 301-311.

Sinclair, H. Number and measurement. In M. F. Rosskopf, L. P. Steffe, & S. Taback (Eds.), *Piagetian cognitive-development research and mathematical education.* Washington, D. C.: National Council of Teachers of Mathematics, 1971.

Skypeck, D. H. The relationship of socio-economic status to the development of conservation of number. Unpublished doctoral dissertation, University of Wisconsin, 1966.

Smedslund, J. The acquisition of transitivity of weight in five- to seven-year-old children. *Journal of Genetic Psychology*, 1963, 102, 245-255. (a)

Smedslund, J. Development of concrete transitivity of length in children. *Child Development*, 1963, 34, 389-405. (b)

Smedslund, J. Patterns of experience and the acquisition of concrete transitivity of weight in eight-year-old children. *Scandinavian Journal of Psychology*, 1963, 4, 251-256. (c)

Smedslund, J. Concrete reasoning: A study of intellectual development. *Monographs of the Society for Research in Child Development*, 1964, 29 (Serial No. 93).

Smedslund, J. Microanalysis of concrete reasoning, I: The difficulty of some combinations of addition and subtraction of the unit. *Scandinavian Journal of Psychology*, 1966, 7, 145-156.

Smock, C. A model for early childhood education. Unpublished manuscript, University of Georgia, 1970.

Smock, C. Discovering psychological principles for mathematics instruction. Paper presented at the Northwestern symposium on cognitive development research and mathematical laboratories, 1973.

Steffe, L. P. The performance of first grade children in four levels of conservation of numerousness on three IQ groups when solving arithmetic addition problems, *Technical Report No. 14, Wisconsin Research and Development Center for Cognitive Learning.* Madison: The University of Wisconsin, 1966.

Steffe, L. P. Thinking about measurement. *The Arithmetic Teacher*, 1971, 18, 332.

Steffe, L. P. Relationships between mathematical and genetic structures: Cardinal and ordinal number. Paper presented at the meeting of the National Council of Teachers of Mathematics, Houston, 1973.

Steffe, L. P., & Johnson, D. C. Problem solving performance of first-grade children. *Journal for Research in Mathematics Education*, 1971, 2, 50-64.

Sullivan, E. V. Piaget and the school curriculum: A critical appraisal. *The Ontario Institute for Studies in Education*, 1967, No. 2.

Uprichard, A. E. The effects of sequence in the acquisition of three set relations: An experiment with preschoolers. *The Arithmetic Teacher*, 1970, 17, 597-604.

Van Engen, H. Foreward. In R. L. Carey & L. P. Steffe, An investigation in learning of equivalence and order relations by four- and five-year-old children. *Research Paper No. 17, Georgia Research and Development Center in Educational Stimulation.* Athens: University of Georgia, 1968.

Van Engen, H. Epistemology, research and instruction. In M. F. Rosskopf, L. P. Steffe, & S. Taback (Eds.), *Piagetian cognitive-development research and mathematical education.* Washington, D. C.: National Council of Teachers of Mathematics, 1971.

Vygotsky, L. S. *Thought and language.* Cambridge: MIT Press, 1962.

Wohlwill, J. F. The abstraction and conceptualization of form, color, and number. *Journal of Experimental Psychology*, 1957, 53, 304-309.

Wohlwill, J. F. A study of the development of the number concept by scalogram analysis. *Journal of Genetic Psychology*, 1960, 67, 345-377.

Wohlwill, J. F. From perception to inference: A dimension of cognitive development. In I. E. Sigel & F. H. Hooper (Eds.), *Logical thinking in children.* New York: Holt, Rinehart & Winston, 1968.

Wohlwill, F. J., & Lowe, R. C. An experimental analysis of the development of the conservation of number. *Child Development*, 1962, 33, 153-167.

Wolf, R., & Kolpfer, L. Program TSSA' test scorer and statistical analysis 2. *Computer Program Library: Social Sciences Division*, Chicago: University of Chicago, 1963.

Zimiles, H. A note on Piaget's concept of conservation. *Child Development*, 1963, 34, 691-695.

Zimiles, H. The development of conservation and differentiation of number. *Monograph of the society for Research in Child Development*, 1966, 31.